2012

Too continuous to continue? Multiple electronic surfaces and derivatives

Aaron Carl West
Iowa State University

Too continuous to continue?
Multiple electronic surfaces and derivatives

by

Aaron Carl West

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Physical Chemistry

Program of Study Committee:
Theresa Windus, Major Professor
Andreja Bakac
Mark Gordon
Monica Lamm
Xueyu Song

Iowa State University

Ames, Iowa

2012

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank my advisor Theresa Windus for her support and patience during my time as a graduate student. I would also like to thank Mike Schmidt and Mark Gordon for teaching me about GAMESS. I would like to thank Kurt Christoffel and Pam Trotter for support during my time as an undergraduate student and for giving me my first research experience and direction in my career choices. I would also like to thank my family and friends for their support and patience over the course of my life. I would like to give a special thanks to my family Alice Woodward and Brandon, Carl, and Sandy West as well as my friends Jedia Plum, Nic Amato, Sean Nedd, George Schoendorf, and Nick Atoms.

CHAPTER 1. GENERAL INTRODUCTION

1. General Overview

Quantum mechanical studies typically fall into two major categories: electronic structure and dynamics. In both branches investigations aim to discern and predict physical events for a variety of atomic- or material-based systems. Electronic structure methods primarily model 'static', physical properties while dynamics studies usually address the temporal evolution of some chemical process. While the electronic structure problem remains vastly difficult, researchers continue to make many advances. In the past, this advancement partially stemmed from experimental needs, which created the demand for novel methods to verify experimental results. However, as experimental methods continue to shift more towards time-based data collection, dynamics methods development should steadily increase and hopefully become more tractable, which in turn will create even more demand for efficient electronic structure methods.

As a result of these drives and current computational limitations, most feasible dynamics methods do not utilize high level, *ab initio* electronic structure methods. For even semiclassical dynamics methods that do involve such high level electronic structure methods, these implementations must sacrifice some numerical accuracy for computational feasibility in both the dynamics and electronic structure aspects. This notion is especially true as chemical system size increases. Then, in all likelihood, most novel chemical observations will come from methods that maintain qualitative rather than quantitative accuracy. For current implementations of *ab initio* treatments, maintaining even this level of accuracy remains difficult except for the smallest of chemical systems. This expense has led to the development of a variety of cheaper but still *ab initio* treatments that include analytical

nuclear derivatives. In the present work, an analytical nuclear derivative implementation is considered for an approximate but less expensive high level method called occupation restricted multiple active space (ORMAS). For the ORMAS method, first-order analytical derivatives are validated over a range of nuclear geometries.

2. Dissertation Organization

The current document includes six chapters, which are published, accepted, or prepared for submission in various journals. Chapter 2 introduces a study on the triplet surface of $O+C_2H_4$. Chapter 3 presents a study on the singlet and triplet surfaces of $O+C_2H_4$ with a focus on biradical species. Chapter 4 continues the previous study with a more extensive study on the energetically lowest-lying singlet surface of $O+C_2H_4$. Chapters 2-4 rely on a several multireference methods to model ground and excited states and on some single-reference methods. Chapter 5 presents a multireference study on $SiCH_4$ and butadiene with ORMAS energy, first-order nuclear derivative, and first-order nuclear derivative coupling contours, which characterize whether or not the ORMAS approximation effects correct analytical derivatives with seemingly smooth energies over a range of geometries. Chapter 6 presents the interface of NEWTON-X (a dynamics driver program) and GAMESS (an electronic structure suite) along with a dynamics study on CNH_4^+ that demonstrates the possible effects of the ORMAS approximation on product distributions.

3. Theoretical Background

Many important discoveries led to the creation of quantum mechanics. Of those important discoveries, two developments tend to stand out: Planck's solution to the black-body radiation problem and Schrödinger's formulation of wave equations.

3.1 Planck

The black-body phenomenon[1,2] occurs when a hollow metal box with a single, closed, pinhole-sized aperture is insulated and heated (via electrical current by electrical leads in the box walls) to equilibrium at some chosen constant temperature (e.g. ~700 ^{o}C). At this point, the insides of the box glow, and quickly opening the aperture releases a spectral sample without initially disturbing the equilibrium. From these samples physicists made two main observations. First, the intensity peak shifts to smaller wavelengths as the temperature increases. Second, the total energy radiated increases as temperature increases. At the time, physicists mostly explained these two observations through thermodynamics. Then, they went on to treat the radiation as standing waves whereby the allowed mode number is proportional to the frequency of the radiation. However, as the observed wavelength approaches zero, the energy from the aperture should become infinite – a property which is not observed in the experiment. The physical lack of infinite energy in the presence of an almost zero wavelength is termed the "Ultraviolet Catastrophe." In order to explain this observation, Max Planck first considered the radiation laws of the day. Planck figured out that the different, separate laws give correct results within the limit of either small or large wavelengths, but at that time no single law produced the black-body radiation spectrum. Planck then literally guessed at the true expression by forming an empirical law. However, in order to explain his result, Planck hypothesized the box walls can only absorb and emit radiation in a discrete rather than continuous fashion and derived the black-body spectral distribution by forming a Maxwell-Boltzmann average with discrete rather than continuous summations. He empirically determined Planck's constant h from the experimental data of his time. In one interpretation Planck called these discrete energy changes "quanta".[1,2]

3.2 Schrödinger Equations

No proof exists for the Schrödinger equations, and such a proof never will exist as long as the equations build off of a physical law (i.e. the conservation of energy). However, a one-dimensional classical wave equation analogy provides some understanding of the Schrödinger hypothesis relative to classical mechanics. The one-dimensional wave equation is a second-order partial differential equation and happens to separate into two ordinary differential equations of the coordinate variable and the time variable. Eq. 1 is the energy conservation law from classical mechanics where E represents energy, p represents momentum, m represents mass, and V represents an arbitrary potential.

$$E = \frac{p^2}{2m} + V \tag{1}$$

From the conservation of energy (Eq. 1), a particle energy, wavelength, and velocity or momentum can be defined. Inserting the particle momentum into the wave equation and symbolizing the coordinate wave amplitude by Ψ yields the one-dimensional time-independent, or coordinate, Schrödinger eigenvalue equation (TISE)[3-11] with coordinates x, Hamiltonian operator H, and total energy E in Eq. 2.

$$H\Psi(x) = E\Psi(x) \tag{2}$$

Through the Hamiltonian operator, the TISE contains analogues for both momentum and coordinate variables: the momentum becomes the $-i\hbar$d/dx operator while the coordinate operator is the multiplicative operator, or potential, V. In general, the wavefunction Ψ can be complex with real and imaginary parts. At the current time, this analogy pervades quantum theory as how to replace classical terms by the corresponding quantum counterparts.

The area of quantum chemistry hinges on so-called stationary state energy eigenvalues E and eigenfunctions Ψ, or wavefunctions, of the TISE. However, later on, Schrödinger included the time-dependent amplitude and derived the one-dimensional time-dependent Schrödinger equation (TDSE)[8] with coordinates x and time t in Eq. 3.

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = H\Psi(x,t) \tag{3}$$

Solving the TDSE produces the total energy eigenvalue and wavefunction in both coordinates and time. In Eq. 3, i equals $\sqrt{-1}$, and $\hbar$ equals Planck's constant h divided by 2π where the denominator constant originates from the periodicity of the general solution for the wave equation.

3.3 Born-Oppenheimer (BO) Approximation

With regards to molecular systems, the TISE contains both electronic and nuclear terms for many electrons and nuclei. In the order of zero-, one-, and two-electron terms, the non-relativistic, molecular, total Hamiltonian H is given in atomic units in Eq. 4.

$$\begin{aligned} H &= \left(-\frac{1}{2}\sum_{\alpha}^{nuclei}\frac{\nabla_\alpha^2}{M_\alpha} + \sum_{\alpha>\beta}^{nuclei}\frac{Q_\alpha Q_\beta}{R_{\alpha\beta}}\right) + \sum_{i}^{electrons}\left(-\frac{1}{2}\nabla_i^2 - \sum_{\alpha}^{nuclei}\frac{Q_\alpha}{r_{i\alpha}}\right) + \sum_{i>j}^{electrons}\frac{1}{r_{ij}} \\ &= \left(T^N + V^N\right) + \left(T^e + V^{eN}\right) + V^e \end{aligned} \tag{4}$$

Equation 4 gives the nuclear kinetic T^N, nuclear potential V^N, electronic kinetic T^e, electron-nuclear potential V^{eN}, and inter-electronic potential V^e operators. Symbols α and β label the nuclei while symbols i and j index the electrons. The ∇_α^2 represents the Laplacian of nucleus α where ∇ is pronounced 'del'. For nucleus $\alpha(\beta)$, the $M_\alpha\left(M_\beta\right)$ stands for the mass while the $Q_\alpha\left(Q_\beta\right)$ denotes the charge. The vector r_{ij} represents the distance between the i^{th} and j^{th} electron, the vector $r_{i\alpha}$ represents the distance between the i^{th} electron and the α^{th}

nucleus, and the vector $R_{\alpha\beta}$ specifies inter-nuclear distances. Electronic motions are often assumed to instantaneously adjust to nuclear motions based on a nucleus-to-electron mass ratio on the order of 1.0 x 10^3 (i.e. ~O(10^3)). Neglecting the above T^N and V^N operators in the total Hamiltonian H results in the non-relativistic, electronic Hamiltonian H^{el}. With the electronic Hamiltonian H^{el} rather than the total Hamiltonian H in Eq. 4, the TISE becomes Eq. 5.

$$H^{el}\Psi(r;R) = E^{el}\Psi(r;R) \tag{5}$$

Solving Eq. 5 results in the multiple electronic solutions $\Psi(r;R)$ where r(R) represents electronic(nuclear) coordinates, and denoted by E^{el}, the resultant electronic energies occur within a specific approximation called the Born-Oppenheimer (BO) approximation.[12] Given stationary nuclei, adding the nuclear repulsion V^N onto the electronic energy E^{el} produces the potential energy surface (PES). Unless otherwise explicitly noted in the present chapter from this point onwards, the term Hamiltonian refers to the non-relativistic, molecular, electronic Hamiltonian with no external fields (e.g. magnetic field).

In order to understand when the BO approximation is not appropriate, solutions to the TISE with total Hamiltonian H must be considered. The term rovibronic denotes these TISE solutions and describes rotational and vibrational states. The total rovibronic solution is constructed from these rovibronic solutions as follows. The rovibronic wavefunction Ψ_{RV}^{Total} can be expanded in arbitrary electronic wavefunctions Ψ_I^e and arbitrary nuclear wavefunctions $\Psi_{K,RV}^I$ (Eq. 6). For the electronic wavefunctions, several approximations can be made at this point. First, restricting the electronic wavefunctions to BO electronic wavefunctions Ψ_I decreases the number of expansion terms (i.e. only a few bound states)

needed to retain some level of accuracy in this ansatz (i.e. assumed form of the rovibronic expansion).[13] Second, if the geometric phase effect[14-18] can be neglected, the BO wavefunctions can be limited to real-valued functions. Third, deriving from the center of mass motion, mass polarization terms[19] are also neglected in the following expansion. Expanding the rovibronic wavefunction Ψ_{RV}^{Total}, substituting this expansion into the TISE with the total Hamiltonian H from Eq. 4, neglecting the additive nuclear potential V^N, and integrating over electronic coordinates r yields the coupled rovibronic equations (Eq. 7).

$$\Psi_{RV}^{Total}(r,R) = \sum_{I,K} \Psi_I^e(r;R)\Psi_{K,RV}^I(R)\beta_{KI} \approx \sum_I \Psi_I(r;R)\Psi_{RV}^I(R) \tag{6}$$

$$\left(T^N + E_I(R) - K^{II}(R) - E^{Total}\right)\Psi_{RV}^I(R) = \sum_{J\neq I}\left(-H^{IJ}(R) + \sum_\alpha\left(\frac{1}{M_\alpha} f_\alpha^{IJ}(R)\frac{\partial}{\partial R_\alpha}\right)\right)\Psi_{RV}^J(R) \tag{7}$$

In Eq. 6, $\Psi_I^e(r;R)$ is the I^{th} arbitrary electronic wavefunction, $\Psi_{K,RV}^I(R)$ is the K^{th} rovibronic wavefunction for the I^{th} electronic wavefunction, β_{KI} is the coefficient for the given expansion term, $\Psi_I(r;R)$ is the I^{th} BO electronic wavefunction, and $\Psi_{RV}^I(R)$ is the overall rovibronic function for the I^{th} BO electronic wavefunction. In Eq. 7, T^N is the previously defined nuclear kinetic operator, E_I is the I^{th} BO electronic energy, E^{Total} is the total energy (minus V^N if neglected above), $\frac{\partial}{\partial R_\alpha}$ is a nuclear derivative (i.e. the del operator ∇_α) for nucleus α, M_α is again the mass for nucleus α, and the additional terms are defined in Eq. 8-10.

$$f_\alpha^{JI}(R) = \left\langle \Psi_J(r;R)\left|\frac{\partial}{\partial R_\alpha}\Psi_I(r;R)\right.\right\rangle_r \tag{8}$$

$$H^{JI}(R) = -\sum_\alpha \frac{1}{2M_\alpha}\left\langle \Psi_J(r;R)\left|\frac{\partial^2}{\partial R_\alpha^2}\Psi_I(r;R)\right.\right\rangle_r \tag{9}$$

$$K^{II}(R) = -\sum_{\alpha} \frac{1}{2M_{\alpha}} \left\langle \frac{\partial}{\partial R_{\alpha}} \Psi_I(r;R) \middle| \frac{\partial}{\partial R_{\alpha}} \Psi_I(r;R) \right\rangle_r \tag{10}$$

In Eq. 8 all the f_{α}^{JI} terms represent the first-order couplings between the Ith and Jth BO electronic states while in Eq. 9 the -H^{JI} contains second-order couplings. The -K^{II} is coined the adiabatic, or the BO diagonal, correction (Eq. 10).

The BO approximation results from the following approximations: neglecting terms from Eq. 8-10 (i.e. a single product essentially appears in Eq. 6) and the zero-order total Hamiltonian is taken to be the electronic Hamiltonian (i.e. no T^{N} and V^{N}). This second approximation neglects nuclear kinetic operator effects on the adiabatic electronic wavefunction (i.e. no -K^{II} term). The BO approximation results in a set of BO wavefunctions, which form an electronically adiabatic basis. The BO wavefunctions are often referred to as adiabatic wavefunctions. However, the adiabatic approximation includes terms from the BO approximation and the adiabatic correction term (i.e. the -K^{II} term) and results in a separate set of wavefunctions. From this point forward, the distinction between BO and adiabatic wavefunctions is suppressed unless explicitly required. After the discussion on how to solve the electronic TISE, the coupled rovibronic equations will prove useful when the BO approximation is inappropriate (i.e. when off-diagonal derivative coupling terms between different adiabatic solutions become non-negligible). In terms of the accuracy of the BO approximation, one last item is worth mentioning. A more recent semi-classical analysis[20] shows the BO energy error to be ~O(m/M)$^{6/4}$ for the electronic mass m and an average-like nuclear mass M.

3.4 Single Reference Wavefunctions

Within the framework of the non-relativistic and BO approximations, many-particle Schrödinger equations can be defined for molecular systems. These electronic TISEs were initially and approximately solved by several mean field theories.[21-25] In a system with n total electrons, a Hartree product (HP)[21-23] Ψ_{HP} wavefunction (Eq. 11) might serve as an initial trial function that approximates the non-relativistic, electronic wavefunction.

$$\Psi_{HP} = \varphi_1(1)\varphi_2(2)...\varphi_n(n) \tag{11}$$

As shown in Eq. 11, a HP consists of a product of n separate functions, and each individual function φ represents a one-electron wavefunction, which is called a spin-orbital. Each spin-orbital consists of spin and spatial parts. (From this point forwards, a spin-orbital is simply referred to as an orbital until further designation.) Despite a convenient product form, HP wavefunctions do not display some very important physical requirements for electrons. First, the exchange of any two electrons should not affect the electronic probability density; for a HP, such an electronic permutation changes the probability density given different functions for different orbitals. Second, experiment supports that electrons are fermion particles, and the antisymmetry principle requires that fermion particle permutations result in an eigenvalue of negative one in Eq. 12. So, for a permutation $\hat{P}_{ij}$ of electrons i and j, an acceptable trial wavefunction Ψ must obey Eq. 12.

$$\hat{P}_{ij}\Psi = -\Psi \tag{12}$$

Functions can typically be split into symmetric and antisymmetric parts, so arbitrary (non-zero and not completely symmetric) functions can be converted into antisymmetric functions.

As a result, electronic wavefunctions Ψ can be formed from approximate wavefunctions $\Psi_{initial}$ by all possible permutations $\hat{P}_i$ on two and more electrons with parity p (Eq. 13).

$$\Psi = \sum_{i=1}^{n!} (-1)^p \hat{P}_i \Psi_{initial} = \hat{A}\Psi_{initial} \tag{13}$$

Eq. 13 defines the antisymmetrizer operator $\hat{A}$, which applies parity p and permutes either electrons among orbitals or orbitals among electrons to produce n! different products. Applying the antisymmetrizer $\hat{A}$ to a HP wavefunction yields Eq. 14 for a Slater determinant ψ_{Slater},[26] which serves as a trial wavefunction for Hartree-Fock (HF)[24,25] wavefunction ψ_{HF}.

$$\psi_{Slater} = |\varphi_1(1)\varphi_2(2)...\varphi_n(n)| \tag{14}$$

In order to find the orbitals that give the best (i.e. lowest energy) solution for a single determinant wavefunction, the Hamiltonian expectation value (i.e. an 'average' energy over many identically prepared, identical systems) of the Slater determinant is formed, and a first-order variation is applied to minimize this energy expectation value. A method such as Lagrangian multipliers keeps the orbitals orthonormal during this variational process and eliminates any need for renormalization of the wavefunction. At the solution a particular resultant Lagrangian multiplier matrix exists. Subsequent diagonalization of this Lagrangian multiplier matrix provides a particular unitary transformation. This unitary transformation produces the canonical orbitals, which are solutions to the effective one-electron Schrödinger equations called the canonical HF equations (Eq. 15-17).

$$f_i\varphi_i = \left[h_i + \sum_i (J_i - K_i)\right]\varphi_i = \varepsilon_i\varphi_i \tag{15}$$

$$J_j\varphi_i(1) = \int \frac{\varphi_j^*(2)\varphi_j(2)}{r_{12}} dr_2\varphi_i(1) \tag{16}$$

$$K_j\varphi_i(1) = \int \frac{\varphi_j^*(2)\varphi_i(2)}{r_{12}} dr_2 \varphi_j(1) \quad (17)$$

The canonical HF equations produce the eigenvalues ε_i, which are called orbital energies. The f_i is called the Fock operator. The h_i is the one-electron (kinetic and electron-nuclear) operator from the third and fourth terms inside the second parentheses in the summation in Eq. 4. J_i is the Coulomb operator, and K_i is the exchange operator, which results in non-zero contributions to the expectation value for only same-spin electronic states. These two operators originate from the two-electron terms in Eq. 4. In addition, application of the one- and two-electron operators and integration over electronic coordinates r produce the corresponding electron integrals. For example, for orthornormal orbitals, Eq. 18 gives the one-electron integrals h_{ii}.

$$h_{ii} = \langle \varphi_i | h_i | \varphi_i \rangle_r \quad (18)$$

The HF equations are a set of complicated integro-differential equations. Through the Coulomb and exchange operators, each effective orbital and its energy depend on the other orbital eigenfunctions. As a result, the HF equations must be solved in an iterative fashion, or in a self-consistent field (SCF). Converting the orbital functions into linear combinations of basis functions is quite preferable in these regards. The linear combination of atomic orbitals (LCAO) method converts these equations into a set of linear algebra equations called the Roothaan-Hall equations[27]. In LCAO, the molecular orbitals (MOs) φ_i are expanded in N atomic orbitals (AOs) χ_p with atomic orbital expansion coefficients c_{pi} (Eq. 19).

$$\varphi_i = \sum_p^N c_{pi}\chi_p \quad (19)$$

The atomic orbital expansion coefficients are often called MO coefficients. Despite the convenience of the LCAO method, using N AO basis functions leads to an increase in the number of one- and two-electron interaction terms. For example, the two-electron terms are now ~$O(N^4)$ where N equals the number of basis functions.

The HF solution typically results from optimized MOs that actually result in the lowest energy expectation value within the single determinant and basis set approximations. (The HF energy does not directly result from the sum of the individual orbital energies.) The resultant HF determinant corresponds to a single configuration with n electrons, and such a so-called single reference solution may very well be qualitatively correct. However, if the physical system has several important Lewis structures, or configurations, the HF solution becomes inappropriate. At some point, any single reference method will inevitably yield qualitatively incorrect physical observables such as bond dissociation energies, which often involve several dominant electron configurations. So, the HF solution typically represents one of a finite number of solutions in a finite atomic basis set expansion.

3.5 Configuration Interaction (CI) Wavefunctions

The energy difference between the exact non-relativistic energy and the HF energy is called the correlation energy. The configuration interaction (CI) approach can be employed to recover the missing energy contributions, but the CI method does not vary the MO coefficients. Instead, in a linear ansatz, a CI wavefunction is expanded as a finite number of spin-adapted determinants, or configuration state functions (CSFs). (From this point onwards, the term orbital now indicates the spatial function of a spin-orbital, which is usually represented as a direct product of space and spin functions.) In contrast to a single reference method, the CI method involves many CSFs and results in ground and excited state energies.

Taking the HF CSF ψ_{HF} to be the first CSF in the expansion only for convenience, Eq. 20 represents the CI wavefunction Ψ_{CI} where different CSFs ψ are formed by occupying different MOs by single excitation S, double excitation D, or higher excitations out of the reference HF CSF ψ_{HF}.

$$\Psi_{CI} = C_{HF}\psi_{HF} + \sum_{S} C_S\psi_S + \sum_{D} C_D\psi_D + \ldots \quad (20)$$

In Eq. 20, the subscripted coefficients C are termed CSF, or CI, coefficients and are variationally optimized. In the case of the full CI wavefunction, the expansion contains all possible Slater determinants from a given basis set: the n electrons are distributed in all possible ways into all the orbitals. Within the BO and basis set approximations, the full CI approach solves the TISE and results in the exact non-relativistic energies. In general, with the addition of novel parameters (e.g. CSF coefficients with linearly independent CSFs) in the linear ansatz and subsequent minimization of all parameters, the bracketing, or Hylleraas-Undheim-MacDonald, theorem[28,29] ensures monotonically lowering, variational, ground and excited state energies.

Because of current computational expense, full CI applications remain limited to the smallest molecules. In fact, for full CI, a linear increase in the number of atomic orbitals results in a factorial increase in the number of CSFs. Computationally feasible CI wavefunctions usually contain much fewer expansion terms than full CI. Any such truncation results in the need for 'excitation schemes' in order to define which expansion terms are included in the wavefunction. In addition, any truncation unfortunately destroys the size consistent and size extensive properties of full CI. In a size consistent method, the energy of a supermolecule with M, non-interacting monomers equals the energy of a single

monomer multiplied by M. In a size extensive method, the energy remains directly proportional to the number of particles as chemical system size increases.

3.6 Multiconfigurational Self-consistent Field (MCSCF) Wavefunctions

Since full CI is computationally expensive, many alternative correlation methods have been developed. One such generally applicable method is the multiconfigurational self-consistent field (MCSCF) method,[30,31] and limiting the CSF expansion makes the MCSCF method tractable. In this method a trial, reference wavefunction is typically expanded in a finite set of orthonormal CSFs and a finite set of AOs to form orthonormal MOs. (At this point, restriction to real orbitals and orbital transformations can be invoked for stationary state, TISE energies.) Exponentiation of any anti-symmetric matrix K provides an orthogonal transformation, which can be normalized and used to vary the MO coefficients. From this matrix K, the unique non-redundant orbital variations, or rotations, κ_{rs} are collected into the supervector $\boldsymbol{\kappa}$. CI variations p_i can also be defined and collected into a supervector p. In this context, the MCSCF method self-consistently optimizes the electronic energy E with respect to non-redundant orbital and CI variations. Eq. 21-22 give the necessary conditions for a critical point of the variational or non-variational MCSCF energy.

$$\frac{\partial E}{\partial \boldsymbol{\kappa}} = 0 \tag{21}$$

$$\frac{\partial E}{\partial p} = 0 \tag{22}$$

Direct optimization of these equations is not feasible in cases with large CI expansions. However, given the exponential parameterization,[32-34] a Taylor expansion up to the second order in the orbital and CI variations about a trial wavefunction yields Eq. 23.

$$E^{(2)}(\kappa,\rho) = E^{(0)} + \frac{1}{2}\begin{pmatrix}\kappa^T & p^T & 1\end{pmatrix}\begin{pmatrix} B & C & w \\ C^T & M & v \\ w^T & v^T & 0\end{pmatrix}\begin{pmatrix}\kappa \\ p \\ 1\end{pmatrix} \tag{23}$$

In Eq. 23, $E^{(2)}$ is a truncated, second-order energy result, and $E^{(0)}$ is the energy of the initial trial wavefunction. The superscript T indicates the transpose of a matrix or vector. The other terms in Eq. 23 represent various first- and second-order derivatives of the energy with respect to orbital and (orthogonal state) CSF variations: w is the orbital gradient, v is the state gradient, B is the orbital Hessian, C is the orbital-state Hessian, and M is the state Hessian. In the last term in the summation in Eq. 23, the upper-left, two-by-two matrix block represents the electronic MCSCF Hessian. Now, optimization can be achieved by the following example implementation: 1) separately minimize CI coefficients with fixed orbitals, 2) apply the critical point conditions to the truncated energy expression, and 3) solve the resultant set of linear equations (with fixed CI coefficients) in order to produce the orbital variations above. At convergence matrix partitioning can be used to verify minimization of a particular electronic state eigenvalue energy with respect to non-redundant, orbital rotations.

As a multireference method, the MCSCF method allows for the possibility of excited state energy optimizations. For MCSCF optimizations of the energetically lowest-lying electronic state of a given spin, orbital optimizations work quite well. However, excited state optimizations tend to result in the lower energy solutions that manifest through state root flipping. A commonly employed solution to this problem is to optimize any energetically lower states with a set of excited state orbitals, and vice versa. In other words, a state-averaged (SA)[35]-MCSCF procedure is used. In formal terms, SA-MCSCF averages the energy expectation values by a set of weights and produces energies for several different

electronic states based on common set of MOs. SA schemes may or may not lead to accurate results relative to the ideal optimization of orbitals with respect to each electronic state or relative to larger active space results.

The most common MCSCF implementation is the full optimized reaction space (FORS)[30,36-38], or complete active space self-consistent field (CASSCF).[31,39] In this implementation, the orbitals are divided up into internal and external sets. The internal set consists of valence orbitals, which can include spectator and active orbitals; the virtual orbitals make up the external orbitals. The internal set contains all occupied orbitals, which include antibonding orbitals with small but non-zero electron occupancies. The spectator set consists of chemical core (e.g. 1s) and selected valence orbitals. For FORS or CASSCF wavefunctions, the truncated CI consists of full CI within the active orbitals rather than all orbitals. In other words, all possible CSFs are generated by distributing active electrons in all active orbitals in all possible ways. However, the factorial increase in the number of CI coefficients with respect to the number of active orbitals limits the number of active orbitals. For smaller molecular systems, computational feasibility allows for either FORS1 or FORS2 treatments.[36-38] In the FORS1 and FORS2 methods, the number of active electrons equals the number of valence electrons, and the number of spectator electrons equals twice the number of chemical core orbitals. For FORS1, the number of active orbitals equals the number of valence orbitals in a minimal valence basis. For FORS2, the number of active orbitals equals the number of active electrons.

A somewhat less common MCSCF implementation is occupation restricted multiple active space (ORMAS).[40,41] ORMAS reduces computational expense by partitioning a total active space into several individually invariant orbital subspaces. Then, defining maximum

and minimum electron occupation values for each subspace results in different excitation schemes. In this *a priori* fashion, the ORMAS approximation produces smaller CI expansions and eliminates so-called 'deadwood'[42,43] determinants to a limited extent with correct use. However, unlike some newer *a priori* techniques[43], the researcher must carefully hand-select the active orbitals in order to obtain the most important determinants in the ORMAS method.

A well-designed MCSCF recovers the majority of quasi-degenerate correlation through a well-chosen active space. Relative to external excitation energies, active orbital excitations require relatively small excitation energies and generally account for the majority of the quasi-degenerate correlation, which should comprise most of the total correlation energy. In contrast, obtaining the majority of dynamical correlation typically requires excitations from the active orbitals into the external orbital set. Such an approach uses the MCSCF approximation as a zero-order wavefunction and further improves this approximation by an additional method (e.g. perturbation theory). Accounting for dynamic correlation frequently allows energies to reach chemical accuracy (i.e. 1-2 kcal/mol).

3.7 Multireference Perturbation Theory (MRPT)

When dynamical correlation energy is especially important, perturbation theory is often used to improve on energetics. In this approach the Hamiltonian H is divided or partitioned into zero- order Hamiltonian H_0 terms and perturbation V terms (Eq. 24).

$$H = H_0 + V \tag{24}$$

If the perturbation is small, then the perturbation corrections will yield chemically accurate energies. When a single reference method cannot yield a qualitatively correct description of a chemical system, multireference perturbation theory (MRPT)[44,45] is used to recover

dynamical correlation and has many formulations. However, because of computational expense, second-order multireference perturbation theory (MRPT2) implementations are most often used. MRPT2 methods are either single- or multi-state methods. A single-state, or single root, MRPT2 method considers perturbation corrections only to a single electronic reference energy while a multi-state method incorporates these corrections into several electronic energy roots of interest.

Single-state, second-order, multireference Møller–Plesset perturbation theory (MRMP2)[46] represents one of several MRPT2 methods. In the MRMP2 method in bra-ket notation, the zero-order Hamiltonian is given in Eq. 25. Here, the expansion terms include the zero-order Ψ_I (from the optimized CAS reference), orthogonal Ψ_K (from additional CAS states), and single CSF Ψ_q (i.e. q) states and corresponding $E_I^{(0)}$, $E_K^{(0)}$, and $E_q^{(0)}$ zero-order energies. Furthermore, single and double excitations from the active to virtual orbitals determine the q states. In Eq. 26, canonical orbital $\varphi_i^{canonical}$ formation consists of separate block-diagonalizations of the standard Fock operator F (one-electron operator analogous to a closed-shell Fock operator) in the inactive, active, and virtual orbital subspaces. The standard Fock operator essentially represents the choice of partitioning in MRMP2. Eq. 27 then gives the corresponding zero-order energies, which are determined by the orbital energies ε_i and the one-electron, spinless, diagonal, density matrix elements γ_{ii}^{λ}. Eq. 28 gives zero-, first-, and second-order MRMP2 energy contributions, which can be informally expressed in terms of electron integrals, orbital energies, and density matrices.

$$H_0 = \left|\Psi_I\right\rangle E_I^{(0)} \left\langle\Psi_I\right| + \sum_{K \neq I} \left|\Psi_K\right\rangle E_K^{(0)} \left\langle\Psi_K\right| + \sum_q \left|\Psi_q\right\rangle E_q^{(0)} \left\langle\Psi_q\right| \tag{25}$$

$$F\varphi_i^{canonical} = \varepsilon_i \varphi_i^{canonical} \tag{26}$$

$$E_\lambda^{(0)} = \sum_i \gamma_{ii}^\lambda \varepsilon_i \quad ; \quad \lambda = I,K,q \tag{27}$$

$$E_I^{(0\to 2)} = \langle \Psi_I | H | \Psi_I \rangle - \sum_q \frac{\left| \langle q | H | \Psi_I \rangle \right|^2}{E_q^{(0)} - E_I^{(0)}} \tag{28}$$

In multi-state, second-order, multiconfigurational quasi-degenerate perturbation theory (MCQDPT2),[47] the zero-order CAS references states now derive from the SA-CAS states rather than one state, and average, spinless, one-electron density matrix elements are now used from those same states. The eventual diagonalization of an effective Hamiltonian yields corrections to several electronic roots of interest.

3.8 Coupled Cluster (CC) Methods

Coupled cluster (CC)[48-50] methods represent another set of methods aimed at the recovery of correlation energy. CC methods can be introduced through arguments about non-interacting versus interacting electron pairs and the coupled cluster expansion,[51,52] which incorporates size-extensivity into a wavefunction by including specific excitations (i.e. disconnected cluster terms) that actually lead to size-extensivity. Unlike the CI method, the CC methods first exponentially expand the linear excitation operator before the various excitation operators act on a reference wavefunction (e.g. HF below) (Eq.29).

$$\Psi_{CC} = \exp(1 + T_1 + T_2 + ...)\psi_{HF} \tag{29}$$

The effect of each excitation operator T is defined in Eq. 30.

$$T_1\psi_{HF} = \sum_S C_S \psi_S; \quad T_2\psi_{HF} = \sum_D C_D \psi_D; \quad ... \tag{30}$$

In Eq. 30, all additional terms are defined as in Eq. 20 except for the coefficients C, which are CC amplitudes. However, truncation at a given excitation level and then application of

the exponential to a reference wavefunction produces the disconnected terms (e.g. resulting from T_2). Size extensivity then naturally occurs. The CC methods thus involve applying nonlinear excitation expansions to some reference wavefunction. However, CC methods then involve the optimization of the exponential form of the wavefunction.

Although many CC-based methods exist, this introduction contains only CC methods relevant to the current work. Based on a HF reference, one CC method with particular success is the completely renormalized coupled-cluster singles, doubles, and non-iterative triples (CR-CC(2,3)).[53,54] However, the CC methods can employ more than just HF reference wavefunctions. Generalized valence bond (GVB)[55-57] wavefunctions have also gained some use as reference wavefunctions. In general, the GVB method is a MCSCF method that involves an optimization with non-orthogonal active orbitals for n electrons. The non-orthogonality makes this method computationally expensive. For this reason, strongly orthogonal GVB (SO-GVB)[58] wavefunctions provide a computationally convenient form of the wavefunction whereby orbitals within each electron pair could remain non-orthogonal while orbitals between electron pairs are forced to be orthogonal. In addition, for singlet states, a less expensive SO-GVB reference wavefunction comes from deleting triplet terms and is called a generalized valence bond-perfect pairing (GVB-PP)[59] singlet wavefunction, for which Hurley first proposed the functional form.

Based on a MCSCF or GVB-type reference, the multireference average quadratic coupled-cluster (MR-AQCC)[60,61] method gives an approximately size extensive, a priori, Davidson-type correction to a multireference wavefunction. The correction neglects greater than connected double excitations and accounts for some disconnected quadratic excitation terms in an average way through the energies of the electron pairs. MR-AQCC defines all

pair energies to be the same, average energy, which is the MR-CISD energy averaged over all electron pairs in the molecule. Relative to other multireference CC precursors to the MR-AQCC method, MR-AQCC provides more accurate corrections for smaller reference wavefunctions.

3.9 Nonadiabatic Coupling Matrix Elements (NACMEs) For MCSCF

Until this point, the primary topic of interest has been the electronic TISE and some methods that lead to its energy solutions. However, gradients and other derivatives must be determined for a variety of reasons (e.g. geometry optimizations, dynamics processes that involve electronic transitions). In particular, nonadiabatic couplings (NACs) or nonadiabatic coupling matrix elements (NACMEs)[62-65] partially determine transitions among pairs of different electronic states. Based on the rovibronic expansion in Eq. 6, the coupled rovibronic equations (Eq. 7) already contain the general first-order NAC f_{α}^{JI} expressions (Eq. 8) in terms of nuclear derivatives and adiabatic wavefunctions. However, for implementation purposes, the NACs must be expressed in terms of the MCSCF and the LCAO implementations. Determining NACs at the MCSCF level traditionally requires the construction of and solution to the coupled perturbed (CP)-MCSCF equations. With limitations on chemical system size, the CP-MCSCF implementation serves as a general and flexible method to analytically evaluate gradients and NACs for ground and excited state PESs.

Deriving a first-order NACME expression requires several steps. Before proceeding, several steps simplify the process. In order to greatly simplify the current CP-MCSCF implementation, a SA-MCSCF procedure is used to produce states with a common set of MOs. This common MO set advantageously ensures the different CI states remain

orthogonal and avoids nonorthogonal overlap terms but with an 'average' MO optimization. Also, re-defining the nuclear derivatives $\frac{\partial}{\partial R_\alpha}$ for the MO coefficients simplifies the CP-MCSCF implementation by converting derivative equations into linear equations (Eq. 31).

$$c_{pi}^{\alpha} = \sum_{j} c_{pj} U_{ji}^{\alpha} \tag{31}$$

In Eq. 31, c_{pi}^{α} are the nuclear derivatives of the MO coefficients, c_{pj} are the MO coefficients, and U_{ji}^{α} are the (orbital) response matrix elements. (The response elements are determined later in step 6.)

The first-order NACME generation steps now follow. First, the SA-MCSCF method produces optimized MO and CI coefficients. Second, expanding the adiabatic wavefunctions into CI coefficients and CSFs in the NAC f_{α}^{JI} expression and applying the derivative product rule produces two contributions: CI $\frac{\partial C^I}{\partial R_\alpha}$ and CSF $\frac{\partial \psi_I}{\partial R_\alpha}$ contributions. Third, taking nuclear derivatives (Eq. 32) of the CI eigenvalue problem and projecting the state J CI vector C^J onto the resulting equation eliminates nuclear derivatives of the CI coefficients and produces the preliminary, first-order CI NACME term ${}^{CI}f_{\alpha}^{JI}$ in Eq. 33.

$$\frac{\partial}{\partial R_\alpha}\left[HC^I - E^I C^I = 0\right] \tag{32}$$

$${}^{CI}f_{\alpha}^{JI}(R) = \left(\Delta E^{JI}\right)^{-1}\left(C^J(R)\right)^T \frac{\partial H}{\partial R_\alpha} C^I(R) \tag{33}$$

In Eq. 32, the E^I denotes the MCSCF energy of the Ith electronic state, and the H denotes the MCSCF Hamiltonian, which is further defined by Hamiltonian matrix elements $H_{\lambda\lambda'}$ in Eq.

34. In Eq. 33, the ΔE^{JI} denotes the energy difference between electronic states I and J, and the derivative Hamiltonian $\frac{\partial H}{\partial R_\alpha}$ is further defined in Eq. 35 by its matrix elements.

$$H_{\lambda\lambda'} = \sum_{i,j}^{MO} h_{ij}\tilde{\gamma}_{ij}^{\lambda\lambda'} + \sum_{i,j,k,l}^{MO} g_{ijkl}\tilde{\Gamma}_{ijkl}^{\lambda\lambda'} \tag{34}$$

In Eq. 34, $h_{ij}\left(g_{ijkl}\right)$ denote the one(two)-electron MO integrals, and $\tilde{\gamma}_{ij}^{\lambda\lambda'}\left(\tilde{\Gamma}_{ij}^{\lambda\lambda'}\right)$ denote the one(two)-particle coupling constants.

$$\frac{\partial H_{\lambda\lambda'}}{\partial R_\alpha} = \sum_{i,j}^{MO} \frac{\partial h_{ij}}{\partial R_\alpha}\tilde{\gamma}_{ij}^{\lambda\lambda'} + \sum_{i,j,k,l}^{MO} \frac{\partial g_{ijkl}}{\partial R_\alpha}\tilde{\Gamma}_{ij}^{\lambda\lambda'} \tag{35}$$

In Eq. 35, $\frac{\partial h_{ij}}{\partial R_\alpha}\left(\frac{\partial g_{ijkl}}{\partial R_\alpha}\right)$ denote the one(two)-electron derivative integral terms, which are further defined in Eq. 36(37).

$$\frac{\partial h_{ij}}{\partial R_\alpha} = h_{ij}^{\alpha} + h_{ij}^{U^\alpha} \tag{36}$$

$$\frac{\partial g_{ijkl}}{\partial R_\alpha} = g_{ijkl}^{\alpha} + g_{ijkl}^{U^\alpha} \tag{37}$$

In Eq. 36(37), $h_{ij}^{\alpha}\left(g_{ijkl}^{\alpha}\right)$ denote nonvarational terms, and $h_{ij}^{U^\alpha}\left(g_{ijkl}^{U^\alpha}\right)$ denote variational contributions in the form of responses where U^α indicates the responses U_{ji}^{α}. In the single reference wavefunction section, expectation values of previously defined one-(two-) electron operators give the one(two)-electron integrals $h_{ij}\left(g_{ijkl}\right)$. However, the LCAO method leads to the form of the one(two)-electron MO integrals $h_{ij}\left(g_{ijkl}\right)$ in Eq. 38(39), which indicates the origin of the nonvariational and variational contributions.

$$h_{ij} = \sum_{p,q}^{AO} c_{pi} h_{pq} c_{qj} \tag{38}$$

$$g_{ijkl} = \sum_{p,q,r,s}^{AO} c_{pi} c_{qj} g_{pqrs} c_{rk} c_{sl} \tag{39}$$

Nuclear derivatives of the MO coefficients c_{pi} in Eq. 38-39 lead to the response $h_{ij}^{U^\alpha}\left(g_{ijkl}^{U^\alpha}\right)$ terms in Eq. 36-37 while nuclear derivatives of the one(two)-electron AO integrals $h_{pq}\left(g_{pqrs}\right)$ lead to the nonvariational one(two)-electron derivative AO integrals $h_{pq}^{\alpha}\left(g_{pqrs}^{\alpha}\right)$ in Eq. 36-37. (Expectation values of one-(two-)electron operators with AOs give the one-(two-)electron AO integrals.) At this point, defining the one(two)-particle transition density matrix elements $\gamma_{ij}^{IJ}\left(\Gamma_{ij}^{IJ}\right)$ in Eq. 40(41) will greatly simplify the form of Eq. 33.

$$\gamma_{ij}^{JI} = \sum_{\lambda,\lambda'} C_{\lambda}^{J} \tilde{\gamma}_{ij}^{\lambda\lambda'} C_{\lambda'}^{I} \tag{40}$$

$$\Gamma_{ijkl}^{JI} = \sum_{\lambda,\lambda'} C_{\lambda}^{J} \tilde{\Gamma}_{ijkl}^{\lambda\lambda'} C_{\lambda'}^{I} \tag{41}$$

Grouping together the response terms in Eq. 36-37 and using the transition density matrices from Eq. 40-41 converts Eq. 33 into Eq. 42.

$${}^{CI} f_{\alpha}^{JI}(R) = \left(\Delta E^{JI}\right)^{-1} \left[\sum_{i,j}^{MO} h_{ij}^{\alpha} \gamma_{ij}^{JI} + \sum_{i,j,k,l}^{MO} g_{ijkl}^{\alpha} \Gamma_{ijkl}^{JI} + \sum_{i,j}^{MO} L_{ij}^{JI} U_{ij}^{\alpha} \right] \tag{42}$$

In Eq. 42, all quantities are previously defined except for L_{ij}^{JI}, which is called the transition Lagrangian for electronic states I and J and results from grouping together response terms. Eq. 43 defines the transition Lagrangian matrix elements L_{ij}^{JI}.

$$L_{mi}^{JI} = \sum_{j}^{MO} h_{mj} \gamma_{ij}^{JI} + 2 \sum_{j,k,l}^{MO} g_{mjkl} \Gamma_{ijkl}^{JI} \tag{43}$$

In Eq. 43, all quantities are previously defined. Converting density matrices in Eq. 42 into the AO basis avoids transforming the derivative AO integrals to the MO basis. These steps essentially complete the CI contributions to the first-order NACMEs.

The CSF $\frac{\partial \psi_I}{\partial R_\alpha}$ contributions to the NAC f_α^{JI} expression are now considered. For the fourth derivation step, the CSF contributions can be re-expressed in terms of the (unfolded) one-particle transition density γ_{ij}^{IJ} for electronic states I and J and the integrals of MOs φ_i with the MO nuclear derivatives $\frac{\partial}{\partial R_\alpha}\varphi_i$ (Eq. 44).

$$^{CSF} f_\alpha^{JI}(R) = \sum_{\lambda,\lambda'} C_\lambda^J \left\langle \psi_\lambda \left| \frac{\partial}{\partial R_\alpha} \psi_{\lambda'} \right\rangle_r C_{\lambda'}^I = \sum_{i,j}^{MO} \gamma_{ij}^{JI} \left\langle \varphi_i \left| \frac{\partial}{\partial R_\alpha} \varphi_j \right\rangle_r \right. \right. \tag{44}$$

In Eq. 44, each CSF ψ_λ is expanded with the antisymmetrizer operator in Eq. 13. Fifth, re-expressing Eq. 44 in terms of MO coefficients c_{pi}, responses, and derivative AO integrals S_{pq}^α and back-transforming densities to the AO basis yields a first-order NACME expression (Eq. 45).

$$f_\alpha^{JI}(R) = \Delta E_{JI}^{-1} \left(\sum_{p,q}^{AO} h_{pq}^\alpha \gamma_{pq}^{JI} + \sum_{p,q,r,s}^{AO} g_{pqrs}^\alpha \Gamma_{pqrs}^{JI} \right) + \sum_{i,j}^{MO} \left[\left(\Delta E_{JI}^{-1} L_{ij}^{JI} + \gamma_{ij}^{JI} \right) U_{ij}^\alpha + \gamma_{ij}^{JI} \sum_{p,q}^{AO} c_{pi} S_{pq}^\alpha c_{qj} \right] \tag{45}$$

In Eq. 45, the AO integrals S_{pq} and derivative AO integrals S_{pq}^α are defined in Eq. 46-47.

$$S_{pq} = \left\langle \chi_p(r;R) \middle| \chi_q(r;R) \right\rangle_r \tag{46}$$

$$S_{pq}^\alpha = \left\langle \chi_p(r;R) \middle| \frac{\partial}{\partial R_\alpha} \chi_q(r;R) \right\rangle_r \tag{47}$$

In Eq. 46-47, χ_p denote the AOs, or atomic basis functions, in the LCAO method. The S_{pq}^α terms arise from nuclear derivatives of the one- and two-electron AO integrals. In addition,

S^{α}_{pq} terms also arise in a constraint equation with the responses U^{α}_{ij}. Given orthonormal MOs, such a relation appears in Eq. 48.

$$\frac{\partial}{\partial R_{\alpha}}\left[\sum_{p,q}^{AO} c_{pi} S_{pq} c_{qj} = \delta_{ij}\right] = U^{\alpha}_{ji} + U^{\alpha}_{ij} + S^{\alpha}_{ij} = 0 \tag{48}$$

In Eq. 48, all items are previously defined except the analogous derivative MO overlap elements S^{α}_{ij} and the discrete Kronecker delta function δ_{ij}, which indicates the MOs are orthonormal. So, these known response contributions are determined, expressed in terms of derivative overlaps S^{α}_{ij}, and separated from the unknown response contributions in the next step.

Sixth, nuclear derivatives of the MCSCF critical point conditions similar to Eq. 21-22 produce the CP-MCSCF linear equations (through repeated use of the chain rule). The left-hand side (LHS) contains the electronic MCSCF Hessian matrix from Eq. 23 and the unknown responses. The right-hand side (RHS) contains (second derivatives in total) nuclear derivatives of the electronic orbital and state gradients also from Eq. 23. In addition to the conditions in Eq. 21-22, SA-CP-MCSCF rather than CP-MCSCF requires derivatives of additional critical point conditions from unique state pairs of SA weights and in turn additional derivative AO integral contributions; this consideration is beyond the scope of the present introduction and is not discussed any further. The sixth step also re-introduces the previously eliminated CI coefficient nuclear derivatives (i.e. CI responses). Then, Handy's turnover rule[66] decreases the number (i.e. 3N-6 for nonlinear molecules) of CP-MCSCF equations where N equals the number of atoms and thus improves on efficiency with less linear equations to solve. After solution of the SA-CP-MCSCF equations, this sixth overall

step ultimately produces the unknown responses, which are subsequently contracted against the above RHS to produce the NACs. These steps summarize the first-order NACME construction for the MCSCF method.

In order to complete the first-order NACME determinations, the SA-CP-MCSCF linear equations must be solved in order to produce the NACs. In many current MCSCF calculations, the CI expansions remain relatively large (i.e. $>10^6$). So, iterative linear equation solvers will be more efficient than directs methods that lead to matrix inversion. Because these equations involve the same Hessian matrix that arises in the SA-MCSCF energy, MCSCF convergence methods should be considered. However, unlike SA-CP-MCSCF problem, the MCSCF problem is nonlinear. As a consequence, MCSCF convergence has a very long and rich history,[67] which is beyond the scope of the present introduction. Nonetheless, several aspects of MCSCF convergence prove especially useful in implementing a solution to the SA-CP-MCSCF equations. Before the MCSCF solver iterations begin, redundant variables are first identified. If a variable (e.g. core-core rotation) is included in the variational parameter space and makes no difference to the solution of an iterative solver step (or the final solution), that variable is redundant. While some redundant variables either help or cause no issues, other redundancies hamper or prevent convergence. Second, redundant variables might be identified before each MCSCF solver iteration (e.g. in Eq. 23, when the state gradient v is non-zero, any orbital gradient w elements below a given threshold give a necessary condition for possible redundancies). If a redundant variable does not get eliminated for some reason, the Hessian matrix can have nearly singular eigenvalues, which can lead to convergence difficulties for unmodified iterative methods. Third, the MCSCF solver (e.g. augmented Hessian method[68]) must account for any negative

eigenvalues in the Hessian. The negative eigenvalues might arise from the partial excited state optimization in SA-MCSCF.

So, in order to solve the SA-CP-MCSCF linear equations, the iterative solver must treat at least the above concerns. For example, if the Hessian contains negative or nearly singular eigenvalues, the (linear) conjugate gradient[69] fails. One possible method is the generalized conjugate residual with inner orthogonalization and deflated restarting (GCRO-DR).[70] The GCRO-DR method is categorized as a one of the Krylov (vector) subspace methods, which tend to suffer from stagnation. Stagnation occurs when the method iterations lower the residual norm by smaller and smaller increments until that norm 'stagnates'. This behavior partly results from required restarts where the traditional Krylov subspace always holds the normalized residual vector as the first vector in the subspace. However, unlike older Krylov subspace methods, GCRO-DR suffers less from stagnation because of its inner-outer scheme and recycled subspaces. The inner scheme minimizes the residual of the linear equation with a modified version of the well-known generalized minimal residual (GMRES)[71] method. The outer iteration involves constructing a deflation vector space. This space restricts the allowed search directions in the algorithm. In other words, it deflates the 'redundant' variations. From the full-dimensional linear equations through matrix-vector products, GCRO-DR constructs projected equations in smaller dimensions, extracts problematic eigenvectors (i.e. negative and nearly zero eigenvalued eigenvectors) from these projected equations, builds orthogonal complement vectors that represent the problematic eigenvectors in the full-dimensional space of the original linear equations, and then finally projects out the redundant search directions in the matrix-vector products of the inner scheme. In addition, GCRO-DR allows for recycling these orthogonal complement vector

spaces between different linear equations. This recycling might reduce the number of iterations required to obtain the NACs. However, eliminating any possible redundancies before the SA-CP-MCSCF will maintain computational feasibility in the convergence process.

3.10 Conical Intersections (C.I.s)

PES intersections often play a prominent role in nonadiabatic dynamics. At these intersections the BO approximation often breaks down and the derivative coupling terms in Eq. 7 become non-negligible. So, locating and characterizing these PES intersections, or surface crossings, often proves useful. However, for some chemical systems, such intersections do not necessarily even exist. The noncrossing rule[15,72-74] affirms the possibility of polyatomic crossings and opposes the existence of diatomic crossings with rare exceptions (e.g. exactly nonrelativistic H_2^+ crossing that disappears with relativistic effects). Instead of such important but subtle "exact Hamiltonian" concerns,[73] the present document focuses on crossings of electronic states with the same spin-space symmetry within the nonrelativistic approximation as described by a finite basis with real-valued functions. The description of such crossings first requires some discussion of an adiabatic versus diabatic basis for the Hamiltonian. An adiabatic basis represents electronic states by energy order while a strictly diabatic basis contains non-zero off-diagonal Hamiltonian elements and zero vectors for the NACs. When electronic states are well separated in energy, the adiabatic basis is very convenient. On the other hand, near PES crossings, adiabatically defined PESs suddenly change physical character (i.e. covalent to ionic), and the NACs become essential to electronic state characterization. Researchers first conceived of the diabatic basis in order to transform the couplings into scalar potential terms and thus simplify the representation of

electronic states near PES crossings. However, such strictly diabatic basis is not possible in polyatomic molecules except in one case: the complete (i.e. infinite) adiabatic basis. Applying the yet-unknown diabatic transformation to an adiabatic basis, setting the NACs equal to 0 vectors, and solving for the diabatic transformation proves this concept.[75]

The previous discussion now allows for a more appropriate description of PES crossings. Given two electronic states of the same spin-space symmetry, a PES intersection requires the following two conditions for (real-valued) Hamiltonian matrix elements: two equal, diagonal elements and two zero, off-diagonal elements. In this case within the 3N-6 internal degrees of freedom where N equals the number of atoms, two-state electronic degeneracy occurs in a (3N-8)-dimensional space, which is called the seam space. The branching space denotes the other two-dimensional space. In this case of same spin-space symmetry, two PESs can generally intersect at some nuclear geometry, or in other words, an "accidental same-symmetry" conical intersection (C.I.) point becomes possible[76,77] (even in a finite basis). In this case along the branching seam, plots of the two state energies then display a double-cone shape.

While the most basic idea of the C.I. seems simple, its location in nuclear coordinate space is non-trivial. As the nuclear degrees of freedom increase, PES scans for C.I.s become unfeasible, which initially delayed the development of efficient C.I. algorithms. Nonetheless, many C.I. search algorithms now exist.[65,78-82] A perturbative algorithm[17,73,83] demonstrates several important concepts about the surrounding electronic surface topology. These concepts greatly simplify the discussion of the C.I. algorithm itself and thus are discussed first. Based on an intersection point of M PESs, the Hamiltonian elements are formed (Eq. 49-51).

$$H_{IJ}^{PP}(R) = \left(C^I(R_x)\right)^T H(R) C^J(R_x) \quad \forall I,J \le M \tag{49}$$

$$H_{IK}^{PQ}(R) = \left(C^I(R_x)\right)^T H(R) C^K(R_x) \quad \forall I \le M; \forall K > M \tag{50}$$

$$H_{KL}^{QQ}(R) = \left(C^K(R_x)\right)^T H(R) C^L(R_x) \quad \forall K,L > M \tag{51}$$

In Eq. 49-51, the M PESs intersect at the nuclear coordinate vector R_x, R represents a different nearby nuclear coordinate vector, the superscript T indicates the transpose of a vector, $C^I(R_x)$ is some CI vector for state I at R_x, and H(R) is the Hamiltonian matrix at R. $H_{IJ}^{PP}(R)$, $H_{IK}^{PQ}(R)$, and $H_{KL}^{QP}(R)$ are the Hamiltonian elements in the new basis. The P space contains CI eigenvectors while the Q space contains CI vectors orthogonal to the eigenvectors in the P space. It is important to note that the Q space still has second-order contributions to the energy. So, a first-order Taylor expansion of the Hamiltonian matrix elements about the intersection of M PESs results in Eq. 52-54, which are the conditions for energy degeneracy for arbitrary first-order changes δR in the nuclear coordinate vector R as defined in Eq. 55.

$$\left(g^{IM}\right)^T \bullet \delta R = \left(\frac{\partial}{\partial R}\Delta E^{IM}\bigg|_{R_x}\right)^T \bullet \delta R = 0 \quad \forall I < M \tag{52}$$

$$\left({}^{num}f^{IJ}\right)^T \bullet \delta R = 0 \quad \forall I < J \le M \tag{53}$$

$${}^{num}f^{IJ} = \Delta E_{IJ}(R) f^{IJ}(R) \tag{54}$$

$$R = R_x - \delta R \tag{55}$$

Throughout Eq. 52-54, the lack of subscript α indicates vectors for all atoms rather than one three-dimensional vector for one atom. The energy difference gradient vector g^{IM} is simply obtained as the difference in energy gradients of states I and M at R_x, and the modified NAC

vector ${}^{num}f^{IJ}$ is defined in Eq. 54 as the NAC vector f^{IJ} numerator (i.e. without the energy difference ΔE_{IJ} denominator) evaluated at a nearby nuclear coordinate R, which is defined in Eq. 55. For a two-state intersection (i.e. M=2), the g^{IM} and f^{IJ} directions define the branching space in the adiabatic basis, and the other directions essentially define the intersection space, which is a (3N-8)-dimensional space. In the resultant seam space, the two PESs remain degenerate through the first-order variation δR. Then, given g^{IM} and f^{IJ} do not equal zero and ΔE_{IJ} equals zero, the two-state intersection is a C.I. Also, these directions directly enter into linear equations for the determining the location of a C.I. The C.I. algorithm facilitates a minimum energy C.I. geometry search in which a Lagrangian function of the energy, the C.I.-defining constraints of any two states, and geometry constraints[65] are minimized through second-order derivatives with respect to nuclear coordinates. Beyond C.I.s, confluences[84-86] exist whereby different C.I.s intersect, which is beyond the scope of the present introduction.

3.11 Tully Surface Hopping (TSH)

Until this point, the primary topic of interest has been the solutions to TISE (Eq. 2 for molecular systems). However, dynamical chemical reaction rates actually result from solving the TDSE (Eq. 3 for molecular systems). Like the TISE solutions to many chemical systems, TDSE solutions also currently require approximations for computational feasibility and can also involve multiple electronic and nuclear solutions; and the subsequent focus continues to remain on multiconfigurational-based, nonadiabatic dynamics methods (rather than adiabatic molecular dynamics within the BO approximation). In one formulation or another for large active spaces and small sized systems (i.e. 4-10 atoms), these types of

approximations typically reduce the nuclear wavefunction representation to nuclear trajectories. Surface hopping (SH)[87-93] methods fall into this category of nonadiabatic, semiclassical dynamics methods; and Tully surface hopping (TSH)[89] currently remains the most well known variant of these SH methods. However, before discussing specifics of the TSH algorithm, deriving TSH from the TDSE formula proves useful in interpretations of TSH itself.

Adding time-dependence to individual nuclear wavefunctions in Eq. 6 gives the total wavefunction $\Psi(r,R,t)$. Without loss of generality, adiabatic electronic wavefunctions are used. Associated with each instantaneous electronic state, each different nuclear wavefunction instantaneously describes the nuclear motion, and vice versa. Substituting the total wavefunction $\Psi(r,R,t)$ into the TDSE (Eq. 2) and using the results essentially already obtained in Eq. 7 (via integration over adiabatic electronic wavefunctions in the same form) yields the coupled partial differential equations in Eq. 56 in atomic units.

$$i\frac{\partial}{\partial t}\Psi_{RV}^{I}(R,t) = \left(T^{N} + \left\{U_{I} \equiv E_{I} + V^{N}\right\} - K^{II}\right)\Psi_{RV}^{I}(R,t) + \sum_{J\neq I}\left(H^{IJ} - \sum_{\alpha}\left(\frac{1}{M_{\alpha}} f_{\alpha}^{IJ}\nabla_{\alpha}\right)\right)\Psi_{RV}^{J}(R,t) \tag{56}$$

In Eq. 56, all terms are already defined in Eq. 7 except time t and U_I, which is defined in Eq. 56. An analogous operation and integration over nuclear coordinates produces a second set of coupled equations. Together, these effective Schrödinger equations remain equivalent to the TDSE. However, the focus remains on how to interpret the SH nuclear motion within the TDSE. Messiah[94] expressed the nuclear wavefunction Ψ_{RV}^{I} (for each electronic wavefunction) as a product of amplitude A_I and phase S_I in Eq. 57.

$$\Psi_{RV}^{I}(R,t) = A_I(R,t)\exp\left[\frac{i}{\hbar}S_I(R,t)\right] \tag{57}$$

Against the traditional separation of complex and real variables, the phase and amplitude are allowed to both be complex.[95] Substituting Eq. 57 into Eq. 56 and separating the PES terms from the derivative coupling terms yields two amplitude-phase coupled equations. The first amplitude-phase coupled equation is Eq. 58.

$$\frac{\partial S_I}{\partial t} + \sum_{\alpha}\frac{1}{2M_\alpha}\left(\nabla_\alpha S_I\right)^2 + U_I = F(h) \tag{58}$$

In Eq. 58, all terms are previously defined except for F(h), which contains Planck's constant h, amplitudes, phases, NACs, and more. Neglecting the F(h) term (which is sometimes argued to be a classical limit regardless of whether or not the amplitudes and phases are allowed to be complex numbers) in Eq. 58 essentially and analogously defines Eq. (59).

$$\dot{R} = \left(\frac{\nabla_1}{M_1},\frac{\nabla_2}{M_2},...,\frac{\nabla_{M_{nuclei}}}{M_{M_{nuclei}}}\right)S_I \tag{59}$$

In Eq. 59, $\dot{R}$ denotes the nuclear velocity vector, each subscripted M denotes each nuclear mass for M_{nuclei} total nuclei, each subscripted ∇ denotes the del operator also for M_{nuclei} total nuclei, and S_I denotes the phase for the nuclear wavefunction for the Ith electronic wavefunction. It is important to note that dots over variables always denote time derivatives. This interpretation greatly simplifies the second amplitude-phase coupled equation in Eq. 60, which resembles the continuity of flux equation.

$$\frac{\partial A_I}{\partial t} + \left(\frac{A_I}{2}\right)\nabla\dot{R} + \nabla A_I \bullet \dot{R} = -\sum_{J\neq I} A_J\left(f^{IJ}\bullet\dot{R}\right)\exp\left[-\frac{i}{\hbar}\int^{t}\left(U_J - U_I\right)dt'\right] \tag{60}$$

By Eq. 60, the amplitudes A_I must correspond to the amplitudes for each electronic state and thus lead to state occupation probabilities. Together, Eq. 60 and the classical limit of Eq. 58

still have no practical solutions because of the non-local cross terms. In other words, these equations couple many trajectories on many electronic states. In order to eliminate this complication, SH methods use the independent trajectory approximation (ITA), which naively reduces Eq. 60 to Eq. 61.

$$\frac{\partial A_I}{\partial t} = -\sum_{J \neq I} A_J \left(f^{IJ} \bullet \dot{R} \right) \exp\left[-\frac{i}{\hbar} \int^{t} \left(U_J - U_I \right) dt' \right] \tag{61}$$

In addition, Eq. 61 essentially represents the more familiar concept: fully coherent propagation of the electronic TDSE. Integration of Eq. 61 yields time-dependent electronic state amplitudes within the ITA; however, directly predicting nuclear geometries with non-negligible NACs prevents analytical solutions (i.e. another "non-locality" of NACs). The "on-the-fly", molecular dynamics algorithms change the nuclear coordinates along a single PES until an electronic transition occurs. Then, in traditional SH algorithms, the nuclear coordinates experience the instantaneous force generated from a single, "occupied" electronic state at each particular time for each independent trajectory in an ensemble of trajectories.

So, with the ITA, different SH algorithms essentially involve different schemes for when to allow electronic transitions. The fewest switches algorithms[89,96] represent some of these schemes. One version of the TSH fewest switches algorithm derivation now follows. Within the ITA, Eq. 62 gives the total wavefunction expansion $\Psi(r,R,t)$ in electronic wavefunctions $\Psi_I(r;R)$, which are assumed to be adiabatic without loss of generality.

$$\Psi(r,R,t) = \sum_I A_I(t) \Psi_I(r;R) \tag{62}$$

In Eq. 62, $A_I(t)$ denotes the amplitude for the Ith electronic state. Inserting Eq. 62 into the electronic TDSE, multiplying by $\Psi_J(r;R)$, and integrating over electronic coordinates r yields Eq. 63.

$$i\hbar\left\{\dot{A}_I(t) \equiv \frac{\partial A_I(t)}{\partial t}\right\} = A_I(t)U_I - i\hbar\sum_J A_J(t)\left(f^{IJ} \bullet \dot{R}\right) \tag{63}$$

Now, two definitions of state probabilities serve well for the derivation. Eq. 64 expresses the occupancy probability P_I for electronic state I as defined by transition probabilities P_{IJ} to all other (electronic) states J.

$$P_I = \sum_J P_{IJ} \tag{64}$$

In terms of trajectories, the probability P_I depends on the fraction of trajectories in state I (i.e. N_I) out of the total trajectories N. The TSH method works under the assumption that the electronic amplitudes A_I determine N_I (Eq. 65) as expressed by the electronic density matrix ρ_{II} in Eq. 66.

$$N_I(t) = \rho_{II}(t)N \tag{65}$$

$$\rho_{II} = A_I^* A_I \tag{66}$$

In Eq. 66, the amplitudes and density both depend on time t, but the labels are suppressed for simplicity. Integrating N_I over time results in Eq. 67.

$$P_I = \int_t^{t+\Delta t} -\frac{\dot{\rho}_{II}}{\rho_{II}}dt \tag{67}$$

In Eq. 67, all terms are previously defined except $\dot{\rho}_{II}$, which is defined by Eq. 68 through time derivatives of Eq. 66. The integrand of Eq. 67 represents the effective, normalized, one-way flux out of state I.

$$\dot{\rho}_{II} = 2\mathrm{Re}\left(A_I^* \dot{A}_I\right) \tag{68}$$

Using Eq. 63, 66, and 68 in Eq. 67 yields Eq. 69.

$$P_I = \int_t^{t+\Delta t} \left[-2 \frac{\mathrm{Re}\left(i\hbar^{-1} A_I^* A_I U_I - \sum_J A_I^* A_J \left(f^{IJ} \bullet \dot{R} \right) \right)}{A_I^* A_I} \right] dt \tag{69}$$

Simplifying Eq. 69 and using Eq. 64 yields the transition probability P_{IJ} (Eq. 70).

$$P_{IJ} = 2 \int_t^{t+\Delta t} \frac{\mathrm{Re}\left(\rho_{IJ} f^{IJ} \bullet \dot{R} \right) dt}{\rho_{II}} \tag{70}$$

The 'instantaneously' occupied state does not change between time intervals. Thus, Eq. 70 becomes Eq. 71.

$$P_{IJ} = \frac{2}{\rho_{II}} \int_t^{t+\Delta t} \mathrm{Re}\left(\rho_{IJ} f^{IJ} \bullet \dot{R} \right) dt \tag{71}$$

In order to prevent transitions into the already occupied state I (i.e. prevent negative probabilities), Eq. 72 defines the effective probability g_{IJ} in an ad hoc manner.

$$g_{IJ} = \max\left(P_{IJ}, 0\right) \tag{72}$$

Eq. 73 finally gives the condition for allowing the nuclear forces on a trajectory to result from a new state K (i.e. a hop from state I to state K occurs).

$$\sum_J^{K-1} g_{IJ} < \xi < \sum_J^{(K-1)+1} g_{IJ} \tag{73}$$

In Eq. 73, ξ denotes a random number from 0 to 1. So, the summation term that first exceeds the random number (if any) indicates a possible electronic transition, which is ultimately based on the NACs, nuclear velocity, and the density matrix elements over time all from an instantaneous force from a single electronic state and nuclear coordinates.

As a consequence of the mixing of quantum probabilities with classical nuclear trajectories, inappropriate transitions sometimes occur for which the new potential energy results in a negative kinetic energy in order to maintain energy conservation. These inaccuracies result from quantum effects such as tunneling and are adjusted for in TSH by several ad hoc methods, which are beyond the scope of this introduction. Despite this and other deficiencies, the TSH method represents one of the first tractable semiclassical methods for large active spaces and small sized chemical systems (i.e. 4-10 atoms).

References:

[1]Planck, M. *Ann. Phys.* **1901**, *4*, 553.

[2]Planck, M. *Ann. Phys.* **1901**, *4*, 564.

[3]Schrödinger, E. *Ann. Phys.* **1926**, *79*, 361.

[4]Schrödinger, E. *Ann. Phys.***1926**, *79*, 489.

[5]Schrödinger, E. *Die Naturwissenschaften* **1926**, *28*, 664.

[6]Schrödinger, E. *Ann. Phys.***1926**, *79*, 734.

[7]Schrödinger, E. *Ann. Phys.***1926**, *80*, 437.

[8]Schrödinger, E. *Ann. Phys.***1926**, *81*, 109.

[9]Schrödinger, E. *Ann. Phys.***1927**, *82*, 257.

[10]Schrödinger, E. *Ann. Phys.***1927**, *82*, 265.

[11]Schrödinger, E. *Ann. Phys.***1927**, *83*, 956.

[12]Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *84*, 457.

[13]Born, M.; Huang, K. *Dynamical Theory of Crystal Lattices*, 1954.

[14]Longuet-Higgins, H. C.; Opik, U.; Pryce, M. L.; Sack, R. A. *Proc. R. Soc. London, Ser. A* **1958**, *244*, 1.

[15]Herzberg, G.; Longuet-Higgins, H. C. *Discuss. Faraday Soc.* **1963**, 77.

[16]Aharonov, Y.; Bohm, B. *Phys. Rev.* **1959**, *115*, 485.

[17]Mead, C. A.; Truhlar, D. G. *J. Chem. Phys.* **1979**, *70*, 2284.

[18]Berry, M. V. *Proc. R. Soc. Lond. Ser. A.* **1984**, *392*, 45.

[19]Sutcliffe, B. T. *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 2321.

[20]Takahashi, S.; Takatsuka, K. *J. Chem. Phys.* **2006**, *124*, 144101-1.

[21]Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 89.

[22]Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 111.

[23]Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 426.

[24]Fock, V. *Z. Phys.* **1930**, *61*.

[25]Fock, V. *Z. Phys.* **1930**, *62*, 795.

[26]Slater, J. C. *Phys. Rev.* **1929**, *34*, 1293.

[27]Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

[28]Hylleraas, E. A.; Undheim, B. *Z. Phys.* **1930**, *65*, 759.

[29]MacDonald, J. K. L. *Phys. Rev.* **1933**, *43*, 830.

[30]Ruedenberg, K.; Sundberg, K. R. *Quantum Sci. FIELD Full Journal Title:* **1976**, 505.

[31]Siegbahn, P.; Heiberg, A.; Roos, B.; Levy, B. *Phys. Scr.* **1980**, *21*, 323.

[32]Dalgaard, E. *Chem. Phys. Lett.* **1979**, *65*, 559.

[33]Yeager, D. L.; Joergensen, P. *J. Chem. Phys.* **1979**, *71*, 755.

[34]Dalgaard, E.; Joergensen, P. *J. Chem. Phys.* **1978**, *69*, 3833.

[35]Docken, K. K.; Hinze, J. *J. Chem. Phys.* **1972**, *57*, 4928.

[36]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41.

[37]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M. *Chem. Phys.* **1982**, *71*, 51.

[38]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 65.

[39]Roos, B. *Adv. Chem. Phys.* **1987**, *69*, 399.

[40]Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364.

[41]Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9377.

[42]Ivanic, J.; Ruedenberg, K. *Theor. Chem. Acc.* **2001**, *106*, 339.

[43]Bytautas, L.; Ruedenberg, K. *Chem. Phys.* **2009**, *356*, 64.

[44]Kozlowski, P. M.; Davidson, E. R. *J. Chem. Phys.* **1994**, *100*, 3672.

[45]Malrieu, J.-P.; Heully, J.-L.; Zaitsevskii, A. *Theor. Chim. Acta* **1995**, *90*, 167.

[46]Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.

[47]Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.

[48]Cizek, J. *J. Chem. Phys.* **1966**, *45*, 4256.

[49]Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697.

[50]Bartlett, R. J.; Musial, M. *Rev. Mod. Phys.* **2007**, *79*, 291.

[51]Sinanoglu, O. *J. Chem. Phys.* **1962**, *36*, 706.

[52]Sinanoglu, O. *J. Chem. Phys.* **1962**, *36*, 3198.

[53]Piecuch, P.; Kucharski, S. A.; Kowalski, K.; Musial, M. *Comput. Phys. Commun.* **2002**, *149*, 71.

[54]Piecuch, P.; Wloch, M. *J. Chem. Phys* **2005**, *123*, 224105/1.

[55]Goddard, W. A., III. *Phys. Rev.* **1967**, *157*, 73.

[56]Goddard, W. A., III. *Phys. Rev.* **1967**, *157*, 81.

[57]Goddard, W. A., III; Dunning, T. H., Jr.; Hunt, W. J.; Hay, P. J. *Accounts Chem. Res.* **1973**, *6*, 368.

[58]Bobrowicz, F. W. Ph.D. Dissertation. California Institute of Technology. **1974**.

[59]Hurley, A. C.; Lennard-Jones, J.; Pople, J. A. *Proc. R. Soc. London, Ser. A* **1953**, *220*, 446.

[60]Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481.

[61]Szalay, P. G.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *103*, 3600.

[62]Tully, J. C. *Mod. Theor. Chem.* **1976**, *2*, 217.

[63]Lengsfield, B. H., III; Yarkony, D. R. *Adv. Chem. Phys.* **1992**, *82*, 1.

[64]Lischka, H.; Dallos, M.; Szalay, P. G.; Yarkony, D. R.; Shepard, R. *J. Chem. Phys.* **2004**, *120*, 7322.

[65]Yarkony, D. R. *J. Phys. Chem.* **1993**, *97*, 4407.

[66]Handy, N. C.; Schaefer, H. F., III. *J. Chem. Phys.* **1984**, *81*, 5031.

[67]Schmidt, M. W.; Gordon, M. S. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233.

[68]Yarkony, D. R. *Chem. Phys. Lett.* **1981**, *77*, 634.

[69]Hestenes, M. R.; Stiefel, E. *J. Res. Nat. Bur. Stand.* **1952**, *49*, 409.

[70]Parks, M. L.; Sturler, E. d.; Mackey, G.; Johnson, D. D.; Maiti, S. *SIAM J. Sci. Comput.* **2006**, *28*, 1651.

[71]Saad, Y.; Schultz, M. H. *SIAM J. Sci. Stat. Comput.* **1986**, *7*, 856.

[72]von Neumann, J.; Wigner, E. *Z. Phys.* **1929**, *30*, 467.

[73]Mead, C. A. *J. Chem. Phys.* **1979**, *70*, 2276.

[74]Longuet-Higgins, H. C. *Proc. R. Soc. London, Ser. A* **1975**, *344*, 147.

[75]Mead, C. A.; Truhlar, D. G. *J. Chem. Phys.* **1982**, *77*, 6090.

[76]Xantheas, S.; Elbert, S. T.; Ruedenberg, K. *J. Chem. Phys.* **1990**, *93*, 7519.

[77]Xantheas, S. S.; Atchity, G. J.; Elbert, S. T.; Ruedenberg, K. *J. Chem. Phys.* **1991**, *94*, 8054.

[78]Farazdel, A.; Dupuis, M. *J. Comput. Chem* **1991**, *12*, 276.

[79]Manaa, M. R.; Yarkony, D. R. *J. Chem. Phys.* **1993**, *99*, 5251.

[80]Bearpark, M. J.; Robb, M. A.; Schlegel, H. B. *Chem. Phys. Lett.* **1994**, *223*, 269.

[81]Levine, B. G.; Coe, J. D.; Martinez, T. J. *J. Phys. Chem. B* **2008**, *112*, 405.

[82]Maeda, S.; Ohno, K.; Morokuma, K. *J. Chem. Theory Comput.* **2010**, *6*, 1538.

[83]Yarkony, D. R. *J. Chem. Phys.* **1994**, *100*, 3639.

[84]Yarkony, D. R. *Theor. Chem. Acc.* **1997**, *98*, 197.

[85]Yarkony, D. R. *J. Chem. Phys.* **1998**, *109*, 7047.

[86]Yarkony, D. R. *J. Phys. Chem. A* **2001**, *105*, 2642.

[87]Tully, J. C.; Preston, R. K. *J. Chem. Phys.* **1971**, *55*, 562.

[88]Preston, R. K.; Tully, J. C. *J. Chem. Phys.* **1971**, *54*, 4297.

[89]Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061.

[90]Sholl, D. S.; Tully, J. C. *J. Chem. Phys.* **1998**, *109*, 7702.

[91]Jasper, A. W.; Stechmann, S. N.; Truhlar, D. G. *J. Chem. Phys.* **2002**, *116*, 5424.

[92]Subotnik, J. E. *J. Phys. Chem. A* **2011**, *115*, 12083.

[93]Subotnik Joseph, E.; Shenvi, N. *J Chem Phys* **2011**, *134*, 024105.

[94]Messiah, A. *Quantum Mechanics*; John Wiley: New York, 1962; Vol. 1.

[95]Tully, J. C. *Faraday Discuss.* **1998**, *110*, 407.

[96]Hammes-Schiffer, S.; Tully, J. C. *J. Chem. Phys.* **1994**, *101*, 4657.

CHAPTER 2. $O(^3P) + C_2H_4$ POTENTIAL ENERGY SURFACE: STUDY AT THE MULTIREFERENCE LEVEL

Aaron C. West, Joshua S. Kretchmer, Bernhard Sellner, Kyoyeon Park, William L. Hase, Hans Lischka, and Theresa L. Windus

Abstract:

The $O(^3P)+C_2H_4$ reaction provides a crucial, initial understanding of hydrocarbon combustion. In this work, the lowest-lying triplet potential energy surface is extensively explored at the multiconfiguration self-consistent field (MCSCF) and MRMP2 levels with a preliminary surface crossing investigation; and in cases where additional dynamical correlation is necessary, MR-AQCC stationary points are also determined. In particular, a careful determination of the active space along the intrinsic reaction pathway is necessary; and in some cases, more than one active space must be explored for computational feasibility. The resulting triplet potential energy surface geometries mostly agree with geometries from methods using single determinant references. However, while the selected multireference methods lead to energetics that agree well, only qualitative agreement was found with the energetics from the single determinant reference methods. Challenges and areas of further exploration are discussed.

Introduction:

The O + ethylene reaction not only serves as an important intermediate in the combustion of most fuels[1] but also provides insight into possible reaction mechanisms involved in the low-Earth orbit erosion (LEO) of space vehicles. However, understanding these reaction schemes is a challenge to both computation and experiment. Several low pressure crossed molecular beam experiments (CMBEs) have been performed for this reaction with only the most recent ones discussed here (the reader is referred to references in these works for early experiments[2-4]). Casavecchia et al. used a low energy, tunable, soft ionization mass spectrometer to reduce the background of dissociative ionization that previous CMBEs experienced and calculated the branching ratios of five competing channels. In particular, this method indirectly characterized several new channels, which include the radical CHO channel.[2] Su et al. used step-scan time-resolved Fourier transform infrared (FTIR) to characterize the formaldehyde channel for reactions of $O(^3P)$ with several alkenes.[3,4] Lee et al. used a CMBE with single-photon ionization to directly observe radical CHO and several other channels.[4]

Under collisional free conditions from 287-2000 K, the most dominant channels comprise the species $CH_2CHO + H$, $CH_3 + CHO$, and $CH_2 + H_2CO$.[5] To pass through these three channels, the reaction path must include the ketocarbene that directly follows from oxygen addition to ethylene. From this ketocarbene triplet biradical, the CH_2CHO and H_2CO channels each involve only one appropriate decomposition pathway on the triplet surface. However, without hyperthermal conditions, the CHO channel faces a large barrier on the triplet surface and will involve an intersystem crossing (ISC) to the singlet surface, a hydrogen shift, and then the appropriate decomposition. Overall product distributions

actually remain pressure independent over a temperature range of 287-2000 K and a pressure range of 0.007-1 atm.[5] In fact, Nguyen et al.[5] have shown that adduct collisional stabilization remains negligible even at 100 atm and ~1500 K where less than 10% triplet adduct stabilization occurs. Thus, temperature primarily controls product distributions under these kinds of conditions. For instance, the product percents of the three channels at 298 K are theoretically and experimentally 40±10, 50±10, and 10±5 while at 2000 K theoretically available values are 18.5, 36.9, and 29.1.[5] Of course, as the temperature varies, all channel product distributions change.

This reaction also presents a theoretical challenge all on its own. Several theoretical studies characterized only static, stationary points[6,7] while other examinations considered important reaction pathways.[8-10] Some lower level studies led to important qualitative reaction schemes.[7,8] However, Nguyen et al.[5] and Schatz et al.[11] more generally examined this potential energy surface. From all the above theoretical studies, major highlights include: 1) the initial hydrogen abstraction channel can only compete with the oxygen addition channel in hyperthermal regions; 2) the oxygen asymmetrically adds to one of ethylene's carbons; and 3) the triplet biradical complex rapidly decomposes via the CH_2CHO + H channel. The Schatz and Nguyen studies provide excellent, general O + ethylene surface explorations; however, these studies have discrepancies. For instance, Schatz et al. had difficulty locating some stationary points with the UB3LYP theory level, and some energy barriers differ between these two studies by up to ~7 kcal/mol. These issues will obviously effect differences in dynamics between the two studies. At low temperatures for this system, large portions of the potential will remain inaccessible, and so these discrepancies will not be as significant. However, under LEO conditions, higher temperatures access more of the

electronic surface. Thus, studies of O + ethylene along with more complex combustion reactions require an accurate surface in hyperthermal regions. While Nguyen examined some of the critical biradicals with multi-configurational methods, most species were examined with the single determinate based methods G3,[12] CBS-QB3,[13] and G2M(CC,MP2)[14] as well as MRCI.[15,16] These methods provided single point energies on B3LYP/6-311++G(3df,2p) optimized geometries. Nguyen et al. averaged the single determinant based results to give average barriers for both energetics and dynamics. Since both surfaces have many reaction pathways with radical character throughout, these single determinate methods may not give proper wavefunctions or barriers in certain cases.

In the O + ethylene reaction, both the lowest-lying triplet and singlet play important roles. However, this study mostly examines the lowest triplet surface; future studies will analogously examine the lowest singlet surface. The important lowest-lying triplet reaction pathways are listed below. Nguyen's study contains pathways 1 - 9, but pathway 10 is not located in that study.

1) $^{\bullet\bullet}O + C_2H_4 \rightarrow {}^{\bullet}CH_2CH_2O^{\bullet}$

2) $^{\bullet\bullet}O + C_2H_4 \rightarrow {}^{\bullet}OH + {}^{\bullet}C_2H_3$

3) $^{\bullet}CH_2CH_2O^{\bullet} \rightarrow {}^{\bullet}CH_2CHO + {}^{\bullet}H$

4) $^{\bullet}CH_2CH_2O^{\bullet} \rightarrow {}^{\bullet\bullet}CH_2 + H_2CO$

5) $^{\bullet}CH_2CH_2O^{\bullet} \rightarrow CH_3{}^{\bullet}CHO^{\bullet}$

6) $^{\bullet}CH_2CH_2O^{\bullet} \rightarrow {}^{\bullet}CH_2{}^{\bullet}CHOH$

7) $CH_3{}^{\bullet}CHO^{\bullet} \rightarrow {}^{\bullet}CH_3 + {}^{\bullet}CHO$

8) $CH_3{}^{\bullet}CHO^{\bullet} \rightarrow CH_3{}^{\bullet}CO + {}^{\bullet}H$

9) $CH_3{}^{\bullet}CHO^{\bullet} \rightarrow {}^{\bullet}H + {}^{\bullet}CH_2CHO$

10) $CH_3•CHO• \rightarrow •CH_2•CHOH$

The rest of the paper is organized as follows. The next section outlines the methods and includes the choice of active spaces. The results and discussion summarize the active space investigation, system energetics, and initial findings on relevant surface crossings. Finally, the conclusion discusses findings relative to Nguyen et al. and summarizes overall results.

Methods:

Results from calculations at the MCSCF and MRMP2 levels in this paper were performed with the GAMESS software.[17] To recover the majority of the static correlation, all calculations were carried out with a complete active space self-consistent field (CASSCF)[18-20] level of theory using the aug-cc-pVTZ[21] basis set (the choice of the aug-cc-pVTZ is justified in the discussion of pathway 1). Because of large complete active space (CAS) sizes and biradical species in this study, the determinant-based method[22] and the full Newton-Raphson (FULLNR) converger with augmented Hessian technique[23-25] were used to provide solutions. Restricted Hartree-Fock (RHF) or Restricted Open shell Hartree-Fock (ROHF) calculations with modified valence orbitals[26] provided good MCSCF starting orbitals in most situations. However, Boys localization[27] provided better starting orbitals in difficult cases. To aid in the convergence of the CASSCF and to prevent convergence to any undesired higher spin states, geometry optimization calculations utilized three configuration interaction (CI) states.

CASSCF stationary point searches for species in the reactions above employed analytical gradients and double differenced, numerical Hessians to characterize the stationary points as first order maxima or minima on the potential energy surface and to obtain zero

point energies (ZPE) and frequencies. At each of the triplet stationary points, 5-10 CI states were calculated to check that the lowest electronic state was attained and to obtain a qualitative picture of the location of the nearest excited states. In addition, 5-10 CI states were also found starting from the singlet wavefunction density at the triplet stationary point geometries to obtain an initial understanding of energetically close-lying states. Since any close-lying, excited state densities may differ from the density of the lowest-lying state of identical spin, a state-averaged CASSCF study could provide a more accurate picture of such states. However, such a study is outside the scope of the current paper.

Intrinsic reaction coordinate (IRC) calculations identified reactants and products associated with each transition state. The second order, Gonzalez-Schlegel (GS2) method[28] provided a robust, steepest descent path in mass-weighted coordinates from the transition state to the minima. Near the transition state, all IRC runs started with a step size of 0.05 bohr*sqrt(amu) for 10 steps. This many initial small steps sufficiently ensured smooth IRCs if the FULLNR converger led to a consistent active space. Then, all IRC runs completed with a step size of 0.15 bohr*sqrt(amu).

Once a reaction path was located, single state, second order Multireference Møller-Plesset perturbation (MRMP2)[29,30] single point energies using the aug-cc-pVTZ basis set were performed along the IRC to recover the majority of the dynamic correlation. In reaction pathways where a second order perturbation theory might not be sufficient to account for the dynamical correlation, stationary points and energetics were also obtained with multireference average quadratic couple cluster (MR-AQCC)[31] using Columbus.[32-34] This method not only recovers more dynamical correlation but also has less sensitivity to the choice of active space. Moreover, this method has the advantage that analytic gradients are

available allowing geometry optimizations at a fully correlated level.[35] Force constants are computed by means of finite differences of energy gradients. Due to the computational expense of these calculations, geometries were determined with an (8,6) active space using the cc-pVDZ[21] basis for H and the cc-pVTZ[21] basis without f functions for C and O (referred to as BS1 in this work). In this particular system, the MR-AQCC results show sensitivity to the basis set; so, several single point energies using different basis sets were used – cc-pVDZ on H and cc-pVTZ on C and O (referred to as BS2) and cc-pVTZ on H and aug-cc-pVTZ on C and O (referred to as BS3).

Of course, the choice of the active space is of critical importance. Several active space sizes were examined for each reaction pathway; Table 1 gives only the most appropriate sizes. However, other somewhat appropriate sizes will be discussed as needed in the sections below. At the moment, MRMP2 calculations larger than a (14,14) active space are not feasible in GAMESS. Hence, this restriction limited the best active space for any reaction pathway. In choosing the CAS, the oxygen 2s (O2s) orbital was always placed in the core. Wherever possible, the oxygen 2p (O2p) orbital was included in the CAS; however, this inclusion quite often resulted in an unphysical, in-out correlation orbital (i.e. an additional mostly unoccupied orbital that resembles an O2p with an additional node as shown in Figure 1). It is very important to understand the reasons for these choices. For this oxygen-containing system, the O2s orbital energy is significantly lower than that of the O2p energy. Hence, the O2s should usually not be important in the quasi-degenerate recovery of the correlation. As well, when the O2s was placed in the core or the active space, it remained doubly occupied and non-degenerate with orbitals required in the active space. Various attempts to place the O2s in the active space generally lead to additional difficulties in

keeping an active space with consistent orbitals; these difficulties parallel the O2p-active space difficulties, which are discussed throughout the paper. Hence, placing both the O2s and O2p in the active space is both unnecessary and problematic.

In all of the reactions, a proper space without in-out correlation was not obtained after several, different, standard attempts. As an example, initial calculations constrained the orbital rotations so that the O2p remained a lone pair in the active space. However, when this constraint was released, the O2p flipped into the core and a different bonding orbital moved from the core to the CAS. As well, other calculations initially placed the O2p in the core without constraints; these runs gave an 'almost desired' active space with the O2p in the core and no in-out correlation in the active space. Then, increasing the space to include a lone O2p in the CAS and constraining the core orbital rotations caused an important antibonding orbital to move to the virtual space and an in-out correlation orbital to move to the CAS. Thus, various cases suggest in-out correlation leads to a lower energy orbital root and makes its elimination from the active space difficult. As a result of these convergence issues, care is required to obtain smooth IRCs. For example, pathway 6 yielded a smooth IRC not only with in-out correlation and three O2p orbitals in the (10,10) CAS but also without in-out correlation and two O2p orbitals in the (8,8) CAS (e.g. one lone O2p orbital was placed in the MCSCF core). However, some cases required the three O2p orbitals and in-out correlation in the active space to obtain a smooth IRC. Furthermore, at some IRC points, the active space would require three antibonding orbitals if a CO π presented itself. Pathways 7 and 8 are examples of this situation, and in these cases the inclusion of orbitals that look like a CO π^*-bond automatically incorporated in-out correlation. In general, any other attempt for no in-out correlation and three O2p orbitals in the active space resulted in

either orbital root changes along the IRC or no existence of the desired root near a geometry that resembled the desired transition state.

Results and Discussion:

Tables 2 and 3 give the overall barrier information for all reactions in the introduction whereby each transition state divides each pathway into reactant/intermediate to transition state (Barrier 1) and transition state to intermediate/product (Barrier 2). Table 2 shows the CASSCF barriers with and without ZPE. Now, since both analytic gradients and Hessians at the MRMP2 level are unavailable in GAMESS, geometry optimizations, IRCs, and ZPEs at this level are not only numerically prohibitive but also can be numerically unfeasible for a single state approach to quasi-degenerate, electronic states (where the perturbation correction can be the most important). In particular, even in cases with symmetric stationary points, the MRMP2 hessian requires asymmetric geometries where close-lying states then have no symmetry constraints. Since MR-AQCC ZPE are very similar to the CASSCF ZPE, ZPE are calculated only at the MCSCF level. However, single point MRMP2 calculations were performed along the IRC (Table 3 shows the barriers). While these barriers are not strictly barriers at the MRMP2 level, they should be representative of them. Single point MRMP2 data at the MCSCF minima and maxima constitute stationary point MRMP2 (SPMRMP2) energy barriers, which always include ZPE. However, as will be shown, single point MRMP2 energies along the IRC result in shifted barriers. These barriers are not true barriers at the MRMP2 level; but they strongly suggest that stationary points at the MRMP2 level do not always coincide with stationary points at the CASSCF level for a given CAS size. For cases where this shift is relatively large and that are critical to the dynamics (reactions 1, 3

and 4), MR-AQCC calculations are performed and discussed below. Table 4 gives the resultant MR-AQCC barriers as compared to CASSCF and shifted MRMP2 results.

In the rest of this section, specifics concerning each of the reactions are presented and discussed. In general, CAS size 1/CAS size 2 defines a CAS pair where the first CAS lacks in-out correlation and the second CAS has in-out correlation; besides the O2p and in-out correlation, the rest of the CAS pair active space is qualitatively the same. Here, comparisons of barriers include ZPE unless otherwise explicitly noted. All additional CAS studies (those not shown in Table 1) have only SPMRMP2, and the description will include what additional bonds were added to the CAS beyond the standard CAS already given in Table 1. Examination of changes in bond length, natural orbital occupation numbers, and localized molecular orbital coefficients over the course of each reaction in different CAS sizes ultimately lead to the choice of the standard CAS.

Pathway 1: $\bullet\bullet O + C_2H_4 \rightarrow \bullet CH_2CH_2O\bullet$

For pathway 1 the standard CAS pair is (6,6)/(8,8). This pathway has A'' symmetry only throughout Barrier 1, which corresponds to the initial oxygen approach. The pathway becomes asymmetric along Barrier 2 far from the transition state at the CASSCF level. As shown in Figure 2, the MRMP2 energy maximum does not coincide with the CASSCF energy maximum; performing single point MRMP2 energies on the CASSCF IRC yields a small, horizontally shifted, MRMP2 energy maximum along the reaction coordinate. As well, this figure also reveals that the MRMP2 significantly changes the relative energetics; and hence the CASSCF Barrier 1 is substantially larger than that for MRMP2. Furthermore, as seen in Tables 2 and 3, neither CASSCF nor SPMRMP2 energies adequately describe Barrier 1. Moreover, SPMRMP2 energies lead to a negative or neutral barrier in the initial

addition of oxygen; this result emphasizes that pathway 1 energy barriers significantly shift with the addition of dynamic correlation. In fact, the reactant minimum also horizontally shifts.

While the different active spaces give quite different CASSCF barriers, the shifted MRMP2 barriers are within 1 and 2 kcal/mol of each other for Barriers 1 and 2, respectively. In addition, the CASSCF ZPE makes less than a 1 kcal/mol difference in the barriers. Taking into account both the transition state and the reactant shifts at the (8,8) CAS, Barrier 1 changes to 1.9 kcal/mol (without ZPE, 2.3 kcal/mol with ZPE) and agrees quite well with the experimental, Arrhenius activation energy of 2.0 kcal/mol.[36,37] As well, Nguyen's average value of 1.3 kcal/mol for Barrier 1 lies within 1 kcal/mol of the experimental activation energy.

In order to justify the size of the basis set selected, single point CASSCF and MRMP2 energies were performed at the aug-cc-pVDZ and aug-cc-pVQZ levels at the aug-cc-pVTZ stationary points of this pathway. As well, CASSCF optimizations and single point MRMP2 energies were performed using the aug-cc-pVDZ basis. All calculations for the basis set validation used the (6,6) CAS for this pathway. In general, the optimized double zeta results differed by 0.2 kcal/mol or less in comparison to the single point double zeta results at the aug-cc-pVTZ stationary points. For Barrier 1 (Barrier 2) with respect to the triple zeta results, the double zeta single point data differed by 0.7 (0.6) and 0.5 (1.1) kcal/mol for the CASSCF and SPMRMP2 data, respectively. For Barrier 1 (Barrier 2) with respect to the triple zeta results, the quadruple zeta single point data differed by 0.2 (0.2) and 0.1 (0.5) kcal/mol for the CASSCF and SPMRMP2 data, respectively. Thus, the triple zeta

results nicely approach the quadruple zeta results even on a pathway where both vertical and horizontal shifting occur.

In order to justify the shifts in the barriers, optimizations for stationary points were carried out at the MRMP2 level. However, as already noted, a state-averaged CASSCF with quasi-degenerate perturbation approach is required here. As a result, slightly shifting the single state, MRMP2 values by the intruder state avoidance (ISA) technique led to pure state minima and a transition state of mixed symmetry for the (6,6) CAS. The two resultant barriers from this ISA MRMP2 optimization match the best shifted barriers to within 1 kcal/mol. In particular, it is important to note that the shifted barriers are obtained using symmetry constraints; furthermore, ISA MRMP2 and shifted transition state geometries closely match. Thus, these shifted barriers are representative of the barriers at the MRMP2 level.

Since the MRMP2 recovery of dynamic correlation may or may not differ between different CAS sizes, we considered another CAS size for pathway 1: (14,14) CAS with a core of O1s, two C1s, O2s, and lone O2p. This active space is comparable to the (6,6) CAS with the addition of all the CH sigma bonds and antibonds to the active space. Barrier 1 for the (6,6) and (14,14) SPMRMP2 results without ZPE do not differ at all (i.e. 0.02 kcal/mol difference in the Barrier 1). However, for the same CAS sizes, the (14,14) SPMRMP2 without ZPE for Barrier 2 gives a 1.9 kcal/mol higher barrier. Only small differences in the geometries of the two active spaces exist; and by performing (6,6) single point MRMP2 energies at the (14,14) stationary points, approximately 1.5 kcal/mol results from changing the size of the active space.

Since this pathway is critical in the overall reaction and dynamic correlation plays a large role in its barriers, MR-AQCC calculations were performed (Table 4). The MRMP2 and MR-AQCC comparisons do not include ZPE since these are very similar for both levels of theory. Using the best basis set results for Barrier 1, the shifted (8,8) and (6,6) MRMP2 result lies within 1 and 1.3 kcal/mol of the MR-AQCC result, respectively. For Barrier 2 the shifted (6,6) MRMP2 results lie within 1 kcal/mol of the MR-AQCC results while the shifted (8,8) now lies 2.8 kcal/mol below the MR-AQCC result. Since having all three O2p orbitals in the active space is important for Barrier1, the shifted (8,8) MRMP2 results agree more with the MR-ACQQ. However, the agreement of the shifted (6,6) with the MR-AQCC for Barrier 2 suggests that in-out correlation may be numerically problematic at the MRMP2 level. It is interesting to note that the (14,14) SPMRMP2 barrier matches the shifted (8,8) MRMP2 Barrier 2 at 19.8 kcal/mol; so, very little is gained by increasing the CAS size.

The above effects of dynamic correlation give several differences in the geometry for the various assignments and levels of theory. The transition state CO σ is 2.05, 1.91, 1.99, 2.13, and 2.17 Å for Nguyen's B3LYP and this study's (6,6) CAS, (8,8) CAS, horizontally shifted (6,6) CAS, and horizontally shifted (8,8) CAS geometries, respectively. Also, the Barrier 2 intermediate OCCH dihedral angle freely rotates from 49 to 63 degrees with an energy change of only 0.02 kcal/mol at the (6,6) CASSCF level. The MR-AQCC and (8,8) CAS geometries do not differ significantly with the exception that the Barrier 2 intermediate has a dihedral angle of 41 degrees.

For this pathway only, Nguyen et al. give constrained optimizations on the CO bond length at the CASSCF(8,8)/cc-pVDZ level with single point CASPT2(8,8)/cc-pVDZ energies (CASPT2/CASSCF) for more accurate barriers in their study. However, it is not clear what

active space was used in this study. For Barrier 2 Nguyen et al. calculated an average value of 25.3 kcal/mol while constrained optimization energies at the CASPT2/CASSCF level gave ~22 kcal/mol, which closely agrees with the MR-AQCC result.

Pathway 2: $\bullet\bullet O + C_2H_4 \rightarrow \bullet OH + \bullet C_2H_3$

The entire reaction pathway has A'' symmetry unlike pathway 1. When oxygen abstracts a hydrogen from ethylene at a typical collisional energy in a CMBE, it encounters a large Barrier 1. As with the case of pathway 1 and many of the other barriers present here, dynamic correlation significantly lowers the barriers for the reaction at the (12,12) CASSCF stationary points (30.3 kcal/mol to 8.9 kcal/mol for Barrier 1 and 12.0 kcal/mol to 4.4 kcal/mol for Barrier 2 from CASSCF to SPMRMP2 values). As a general trend, MRMP2 always energetically lowers CASSCF barriers by a much more substantial amount in the cases of bimolecular products to transition state barriers than in the cases of unimolecular intermediate to transition state barriers.

However, for this case dynamic correlation gives no significant differences in the geometries since the barriers do not horizontally shift by much. The energy change between the SPMRMP2 barriers and the shifted barriers is small (less than 1 kcal/mol) with or without shifts in the minima. However, the geometric changes due to the transition state horizontal shift along the IRC remain small in the (12,12) CAS in comparison to the minima horizontal shifts. CASSCF ZPE changes Barrier 1 by 4.1 kcal/mol – a significant change – and Barrier 2 by 1.5 kcal/mol. For the shifted MRMP2 barriers, Barrier 1 lies 1.4 kcal/mol below Nguyen's average Barrier 1 while Barrier 2 lies 1.8 kcal/mol above Nguyen's average Barrier 2. However, the products in our work have a potential well when the OH is ~3 Å from the closest carbon. At the 50 Å distance, Barrier 2 is smaller by ~1.4 kcal/mol based on the

SPMRMP2 value at the ~3 Å distance. SPMRMP2 Barrier 1 at 50 Å is 0.3 kcal/mol smaller than SPMRMP2 Barrier 1 from the ~3 Å minimum.

CASSCF runs could not produce a (14,14) CAS with in-out correlation for this pathway; the O2p orbital tended to flip from the CAS to the core towards the reactant end of the IRC. We also investigated (8,8) and (10,10) CASs where all non-action CH σs (i.e. those not directly involved in the hydrogen abstraction) from the standard (12,12) CAS were placed in the core and the CC σ was placed in the active space for the (10,10). Both these CASs have in-out correlation. The (8,8) and (10,10) CASs have all SPMRMP2 barriers within 1 kcal/mol of each other; however, these nondegenerate, in-out correlation barriers differ from the (12,12) CAS barriers without in-out correlation by ~1 kcal/mol and ~2.4 kcal/mol for Barriers 1 and 2, respectively. The barriers calculated from the (12,12) CAS represent the overall process better.

Pathway 3: $\bullet CH_2CH_2O\bullet \rightarrow \bullet CH_2CHO + \bullet H$

The standard CAS pair for pathway 3 is (8,8)/(10,10). As with the rest of the pathways on the lowest-lying triplet surface in this paper, pathway 3 is asymmetric. Including the dynamic correlation here gives a large horizontal shift in the transition state but no significant horizontal shift for the products of Barrier 2. However, the shift in the transition state does change the geometries, and hence barrier energies change as well. Between the various levels of theory and shifted geometries, the dissociating H-C distance in the transition state (the action CH σ) is 1.77, 1.62, 1.57, 1.74, and 1.57 Å for Nguyen's B3LYP and this study's (8,8) CAS, (10,10) CAS, horizontally shifted (8,8) CAS, and horizontally shifted (10,10) CAS geometries, respectively. For the MRMP2 barriers with a shifted transition state with respect to the SPMRMP2 barriers, Barriers 1 and 2 both increase

by only 0.7 kcal/mol in the (8,8) CAS while both Barriers 1 and 2 increase by 1.9-2.0 kcal/mol in the (10,10) CAS. Again, minima shifts do not make much difference in this analysis. Furthermore, the shifted (8,8) and (10,10) MRMP2 Barrier 2s are almost identical while the shifted (10,10) MRMP2 Barrier 1 lies 1.6 kcal/mol above the corresponding shifted (8,8) value. CASSCF ZPE lowers Barrier 1 by ~4.0 kcal/mol (again, a rather significant effect) and raises Barrier 2 by 1-2 kcal/mol. Now, both the shifted (8,8) and (10,10) MRMP2 barrier values lie within 1 kcal/mol of Nguyen's average barriers. All of the shifted MRMP2 barriers are within 2 kcal/mol of the MR-AQCC/BS3 results with the (10,10) within 1 kcal/mol for Barrier 1.

For this pathway, including the CC σ in the active space has a small effect since the CC σ length changes by only 0.05 Å. For the SPMRMP2 barriers, incorporating the CC σ into larger (10,10) and (12,12) CAS calculations leads to a ~1 kcal/mol and ~0.5 kcal/mol change from the (8,8) and (10,10) Barriers 1 and 2, respectively.

Pathway 4: $\bullet CH_2CH_2O\bullet \rightarrow \bullet\bullet CH_2 + H_2CO$

The (6,6)/(8,8) CAS pair describes how the main biradical intermediate splits into biradical methylene and formaldehyde. For this reaction dynamic correlation shifts both the transition state and the Barrier 2 minimum. Such shifts substantially alter both the reaction coordinate and the energy of the transition state but does not significantly change the Barrier 2 minimum energy. As a result, the various levels of theory give differences in the geometries. For the transition state, the action CC σ is 2.28, 2.00, 1.94, 2.23 and 2.21 Å for Nguyen's B3LYP and this study's (6,6) CAS, (8,8) CAS, horizontally shifted (6,6) CAS, and horizontally shifted (8,8) CAS geometries, respectively. In addition, Barrier 2 bimolecular intermediate (approaching products) has (6,6) CAS, (8,8) CAS, horizontally shifted (6,6)

CAS, and horizontally shifted (8,8) CAS geometries with CC distance of 4.14, 4.16, 3.40, and 3.50 Å, respectively. This distance results from a hydrogen bonding potential well from the bimolecular intermediate, which approaches the products at larger distances. These geometrical differences result in various energy changes. For the MRMP2 barriers with a shifted transition state with respect to the SPMRMP2 barriers, Barriers 1 and 2 both increase by 3.1 and 5.2 kcal/mol in the (6,6) and (8,8) CAS, respectively. Once again, including the minima shift does not make much difference to these observations. So, even though a significant geometrical change occurs, very little difference in the energetics results. The shifted (8,8) and (10,10) MRMP2 barriers lie within 1 kcal/mol of each other for Barrier 2 while for Barrier 1 the shifted (10,10) value lies 1.9 kcal/mol above the corresponding (8,8) value. For Barrier 1 the SPMRMP2 (6,6) and (8,8) barriers lie within 1 kcal/mol whereas the SPMRMP2 (8,8) Barrier 2 lies 1.3 kcal/mol above the corresponding (6,6) Barrier 2. The MR-AQCC calculations show agreement with the shifted MRMP2 to within 1 kcal/mol for the (8,8) active space. CASSCF ZPE changes all barriers by ~2.2 kcal/mol. So, in comparison to Nguyen's values, the (6,6) MRMP2 Barrier 1 lies 1.9 kcal/mol below the average Barrier1; and the (8,8) MRMP2 Barrier 1 matches Nguyen's average Barrier 1. For the MRMP2 Barrier 2, the (6,6) CAS lies 1.5 kcal/mol above while the (8,8) CAS lies 2.3 kcal/mol above Nguyen's average.

Larger (10,10) and (12,12) CAS that include additional correlation from the formaldehyde two CH σs were investigated. Here, Barrier 2 values are within 1 kcal/mol; however, Barrier 1 values are ~2-3 kcal/mol higher. However, the values compared are SPMRMP2 values, and the shift will be different for different CAS sizes. It should be noted that unlike in pathway 1, the CH σs on different carbons are never degenerate; however, not

including some quasi-degenerate CH σs in the active space might lead to undesirable results as well.

Pathway 5: $\bullet CH_2CH_2O\bullet \rightarrow CH_3\bullet CHO\bullet$

In this pathway a hydrogen shifts from one carbon to the other carbon. As a result of the large (14,14) CAS size, dynamical correlation did not lead to any reportable geometry shifts. In addition, the (14,14) CAS geometry does not differ from Nguyen's B3LYP in any significant way. Also as a result of the large CAS size, the MRMP2 barriers do not exhibit any difference from the SPMRMP2 barriers. No comparison for in-out correlation exists here for various reasons (see below). The CASSCF ZPE give a substantial change of 2.3 kcal/mol for Barrier 1 and 3.4 kcal/mol for Barrier 2. The MRMP2 Barrier 1 lies 1.6 kcal/mol above Nguyen's average Barrier 1; and the MRMP2 Barrier 2 lies within 1 kcal/mol of Nguyen's Barrier 2.

This hydrogen shift between the two carbons could only be explored without in-out correlation; any attempt to find a transition state with in-out correlation in a smaller CAS where the active space does not include degenerate hydrogens failed. Here, in-out correlation orbitals do not lead to a qualitatively correct, ground state wave function. In addition, the (14,14) CAS includes the CO σ in the active space; an attempt to include the O2s, CO σ, and O2p in the core in a (12,12) CAS leads to mixing of the O2s and CO σ. Often times, such a mixing, or polarization, of the O2s leads to improper orbital root changes later on in the IRC. As well, the active space also requires inclusion of the CC σ since this bond is mixed with the shifting hydrogen. Hence, these issues lead to the choice of active space.

SPMRMP2 barriers for an (8,8) CAS with a core of O2s, all non-action CH σs, and O2p, (12,12) CAS with a core of O2s, CO σ, and O2p, and (12,12) CAS with a core of O2s, CH σ nearest the oxygen, and O2p were also found. Barrier 1 values are all within ~1 kcal/mol of each other; however, Barrier 2 values differ by the active space chosen and range from 37.9 to 41.7 kcal/mol. In particular, the (12,12) barriers with the quasi-degenerate hydrogen in the core lies within 1 kcal/mol of the (14,14) barriers.

Pathway 6: $\bullet CH_2CH_2O\bullet \rightarrow \bullet CH_2\bullet CHOH$

In this reaction the hydrogen shifts from carbon to oxygen and requires a (8,8)/(10,10) CAS pair. The core includes the CC σ, which changes by only 0.02 Å over the entire reaction. Here, dynamic correlation does not horizontally shift any of the CASSCF stationary points by a significant amount. However, the Barrier 2 intermediate has two conformers with no geometrical differences besides a hydroxyl rotation. The lowest energy conformer has a HCOH dihedral of 179 degrees (trans-like) while the higher energy conformer has a corresponding 36 degree dihedral angle (cis-like); both conformers have COH angles of 108 degrees with an energy difference of 0.2-0.3 kcal/mol at the CASSCF level. For the shifted MRMP2 barriers with respect to the SPMRMP2 barriers, all barriers change by at most 0.1 kcal/mol. The shifted (10,10) MRMP2 Barrier 1 lies 1.1 kcal/mol above the corresponding shifted (8,8) Barrier 1; and for Barrier 2, the shifted (10,10) barrier lies 1.6 kcal/mol above the corresponding (8,8) value. CASSCF ZPE significantly changes Barrier 1 by ~2.7 kcal/mol and Barrier 2 by ~3.1 kcal/mol both with and without in-out correlation. For the shifted (10,10) MRMP2 barriers, Barrier 1 lies 2.0 kcal/mol lower than Nguyen's average Barrier 1, and Barrier 2 lies 2.8 kcal/mol below Nguyen's average Barrier

2. The smaller (8,8) comparisons to Nguyen's barriers give Barrier 1 lower by 1.1 kcal/mol and Barrier 2 lower by 1.6 kcal/mol.

Pathway 7: CH_3•CHO• → •CH3 + •CHO

This dissociation involves all CO π bonds and must contain in-out correlation to obtain a CASSCF solution with the FULLNR converger. For this pathway dynamic correlation leads to a large horizontal shift in both the transition state and Barrier 2 products. MR-AQCC calculations were not performed here since the previous results show the error at the MRMP2 level is at worst 3 kcal/mol and most likely less. Geometrical differences include the transition state CC σ of 2.12, 1.89, and 2.12 Å for Nguyen's B3LYP, (8,8) CAS, and shifted (8,8) CAS geometries, respectively. In addition, Barrier 2 products have CC σ of 4.09 and 3.40 Å for the CAS and shifted CAS geometries, respectively. For the shift in the barriers from transition state shift as compared to SPMRMP2 barriers, Barriers 1 and 2 increase by ~4 kcal/mol for the shifted barriers – a large effect. As for the effect of the Barrier 2 minimum geometry shift on the energies, this shift increases the overall Barrier 2 by 0.4 kcal/mol. CASSCF ZPE decreases Barrier 1 by 2.3 kcal/mol and increases Barrier 2 by 4.2 kcal/mol. For the shifted MRMP2 barriers, Barrier 1 agrees with Nguyen's Barrier 1, and Barrier 2 is 2.3 kcal/mol higher than the Nguyen values.

A (10,10) CAS with additional correlation from the CH σ nearest the oxygen gives SPMRMP2 barriers within 1 kcal/mol of the corresponding (8,8) barriers. Any attempt to exclude in-out correlation via a core, lone O2p orbital leads to unwieldy iterations towards the dissociation end of the IRC; these large iterations indicate inappropriate orbital root changes.

Pathway 8: CH_3•CHO• → CH_3•CO + •H

This dissociation also involves all CO π bonds and must contain in-out correlation to converge with the FULLNR converger to a CASSCF solution; hence, it requires an analysis almost identical to that of pathway 7. Both the transition state and Barrier 2 products exhibit large horizontal shifts in the IRC with MRMP2. As for the resulting geometry changes, the transition state action CH distance is 1.78, 1.51, and 1.72 Å for Nguyen's B3LYP, (8,8) CAS, and shifted (8,8) CAS geometries, respectively. As with pathway 7, the B3LYP results agree with the shifted CAS geometries in the transition state. In addition, the Barrier 2 product action CH σ is 5.35 and 3.58 Å for CAS and shifted CAS geometries, respectively. There is a dramatic change between the transition state, shifted MRMP2 barriers and the SPMRMP2 barriers; both Barriers 1 and 2 shift by 3.7 kcal/mol. As for the effect of the Barrier 2 minimum geometry shift on the energies, this shift increases the overall Barrier 2 by 0.4 kcal/mol. CASSCF ZPE increases Barrier 1 by 4.3 kcal/mol and Barrier 2 by 1.8 kcal/mol. For the shifted MRMP2 barriers, Barrier 1 lies 2.7 kcal/mol above Nguyen's average Barrier 1 while Barrier 2 lies about 1 kcal/mol higher than Nguyen's average Barrier 2.

A (10,10) CAS with additional correlation from the CC σ gives SPMRMP2 barriers within 1 kcal/mol of the corresponding (8,8) barriers. Again, any attempt to exclude in-out correlation via a core, lone O2p orbital leads to unwieldy iterations towards the dissociation end of the IRC.

Pathway 9: $CH_3\bullet CHO\bullet \rightarrow \bullet H + \bullet CH_2CHO$

In this pathway, hydrogen dissociates and actually requires three O2p orbitals in the active space for the possibility of a smooth IRC. The (12,12) CAS – which does not include all three O2p orbitals – does not have a smooth IRC because a discontinuity appears a few

steps toward the Barrier 2 product where a lone, core O2p violently changes orientation. Likewise, the biradical, active space O2p violently changes orientation at the same IRC point. Thus, the active space must contain all three O2p orbitals in order to produce a smooth IRC. However, for this pathway in a (14,14) CAS, having three active space O2p orbitals and in-out correlation does not lead to a transition state. Thus, in a similar fashion to the in-out correlation problem in pathway 5, the (14,14) CAS here does not lead to a wavefunction which is qualitatively consistent with smaller CAS wavefunctions that lack in-out correlation. However, a (14,13) CAS with no in-out correlation leads to an IRC that consistently maintains its active space towards the Barrier 1 intermediate. However, the first IRC step towards the products leads to unwieldy energy iterations and eventually to too many constrained optimizations in the IRC itself. The unwieldy iterations occur because the converger leads to orbital rotations that try to incorporate in-out correlation as the reaction moves towards the products. In other reaction pathways that include in-out correlation (when possible) from the start, the in-out correlation tends to become or make up the CO π^*. However, with a (14,13) CAS, this reaction pathway starts with no in-out correlation at the transition state and then subsequently develops in-out correlation in a linear combination with the CH σ^* and displaces the CC σ^* during the geometry changes. So, we do not report (14,13) results for Barrier 2.

For this pathway neither the large (12,12) or the (14,13) CAS have much of a horizontal shift in the IRC from dynamic correlation. The minima do not horizontally shift at all from the MRMP2. No noteworthy geometry differences occurred between this study and Nguyen's results. For the SPMRMP2 barriers with respect to the shifted barriers, the (12,12) barriers shift by less than 1 kcal/mol while the (14,13) Barrier 1 does not shift at all.

CASSCF ZPE lowers Barrier 1 by ~5.5 kcal/mol and increases Barrier 2 by ~1.2 kcal/mol. For the shifted MRMP2 barriers, both (12,12) CAS Barrier 1 and 2 correspondingly lie within 1 kcal/mol of Nguyen's values. With the additional O2p in the active space (14,13), Barrier 1 decreases by 2.5 kcal/mol.

Pathway 10: CH_3•CHO• → •CH_2•CHOH

This hydrogen shift requires at minimum a (12,12) CAS. This large CAS size leads to no horizontal shifts. As before with the other CAS sizes, no substantial geometry differences exist in the minima although the dynamic correlation does change the overall barriers. Since Nguyen et al. do not solve for this transition state, no comparisons can be made in this respect. As well, from previous data and the necessary CAS size, the (12,12) should provide an adequate solution for this particular reaction. The CASSCF ZPE decreases Barrier 1 by 3.4 kcal/mol and increases Barrier 2 by 2.5 kcal/mol.

Surface Crossings: Preliminary Investigation at the MCSCF Level

Several factors come into play in the location of relevant, close-lying electronic states. First and foremost, active space choice affects the electronic states found; a reasonable active space will lead to the states of interest. Second, the character (i.e. the spatial and spin characteristics) of an initial guess relative to the desired molecular orbitals along with the number of CI roots requested determines the kinds of states located; a good guess at the initial molecular orbitals along with a reasonable number of CI roots will minimize the number of missed electronic states. Determining a reasonable number of states to request requires some experimentation; the lowest-lying energies from calculations with different numbers of CI roots can always be compared. In this discussion, a resultant

electronic state with an exact density refers to any state of interest obtained by optimizing a particular set of orbitals and CI coefficients under given spin and spatial (if any) symmetries.

On the basis of the above description, CI energy runs with several roots and with both singlet and triplet multiplicities were performed on all the lowest-lying triplet stationary points in order to estimate the relative location of nearby states. In addition, separate energy calculations of the A' surface were run for pathways 1 and 2 to account for higher symmetry states. While certainly not quantitative, these calculations provide a preliminary analysis on the locations of surface crossings, which will be explored in more detail in future work. For this analysis we report any states near the lowest-lying triplet within ~10 kcal/mol and also note any singlet states below the lowest-lying triplet. In the analysis below, ground state means the lowest state of either the triplet or singlet states. Also, energetics for the lowest singlet and triplet are always determined from the exact densities. For any close-lying, higher, inexact roots, we note which exact density (singlet or triplet along with spatial symmetry if any) yields, or derives, the higher root of interest. Furthermore, any barriers reported are SPMRMP2 barriers.

For pathway 1 the transition state and Barrier 2 intermediate have close-lying states. For the transition state, neither the (6,6) nor the (8,8) CAS can adequately describe the triplet A' with an exact density (i.e. orbital root flipping occurs). So, a (4,4) CAS A' energy on the triplet (4,4) CAS A'' transition state and a (4,4) CAS A' transition state were found; this (4,4) CAS contains only the biradical and CO σ/CC π. These calculations show an A'-A'' energy difference of only 0.6 kcal/mol, and the A' transition state also lies 0.6 kcal/mol above the triplet A'' stationary state. These results lie within 1 kcal/mol of other calculations in literature.[5,11] The Barrier 2 intermediate has a triplet state above its singlet ground state by

0.6 and 1.3 kcal/mol for the (6,6) and (8,8) CAS, respectively. In addition, another singlet state derived from the triplet density lies 2.6 kcal/mol above the triplet in both the (6,6) and (8,8) CAS.

For pathway 2, as the products are separating and each going to doublet states, the overall singlet and triplet surfaces for the reactions become degenerate. It is interesting to note that Nguyen et al. found a triplet A' state less than 1 kcal/mol above the triplet A'' ground state; however, the (12,12) CAS did not give this result. It gives the triplet A' state as ~70 kcal/mol above the lowest-lying triplet A'' state with exact densities. However, the previously mentioned (14,14) CAS with root flipping gives a triplet A' state ~10 kcal/mol above the triplet A'' ground state.

Furthermore, SPMRMP2 on this triplet A'' stationary point yields the triplet A' as lying ~0.5 kcal/mol below the triplet A'' state. Although the order of states interchanges with the application of MRMP2, a quasi-degenerate perturbation calculation should give the same ordering as the CASSCF results. However, such a calculation is currently prohibitive in GAMESS in the configuration state function code. This particular example shows the importance of selecting the active space to yield the desired electronic states.

For pathway 3 the Barrier 1 minimum has a singlet state 1.5 kcal/mol above the triplet ground state for both the (8,8) and (10,10) CAS. This minimum corresponds to the Barrier 2 minimum in pathway 1. Examination of the two structures shows that the IRCs for different pathways lead to different conformers, which interchange state orders near the regions of the ISC. This behavior results from the free OCCH dihedral angle (as discussed above in pathway 1 and shows this particular biradical definitely has a surface crossing and possibly several surface crossings at that. In addition to these results, the transition state has a singlet

that exists above the triplet ground state by 7.4 and 3.4 kcal/mol for the (8,8) and (10,10) CAS, respectively. SPMRMP2 corrections to these energy differences in the singlet versus the ground state triplet lead to 5.2 and 3.9 kcal/mol for the (8,8) and (10,10) CAS, respectively. As for Barrier 2, the IRC leads to a doublet whereby the singlet and triplet energies become degenerate for both the (8,8) and (10,10) CAS.

Barrier 1 in pathway 4 corresponds to the predominant triplet-singlet surface crossing biradical already mentioned several times; this biradical will not be mentioned again since the same results hold for all CAS sizes. This pathway contains no other close-lying states based on this analysis; and the lowest-lying triplet state appears to remain the ground state through the rest of the pathway.

In pathway 5 Barrier 2 intermediate has a singlet ground state that lies 60.6 kcal/mol below the lowest-lying triplet state. Furthermore, a second excited singlet state derived from the triplet density lies ~3.6 kcal/mol above the lowest-lying triplet state. As a result, the transition state geometry also has a singlet ground state with a triplet that lies 50.9 kcal/mol above it. The second excited state singlet was not found from the triplet state density suggesting that the ground state singlet density better describes the singlet excited state.

The transition state for pathway 6 has a singlet above the ground state triplet by 0.4-0.5 kcal/mol for the (10,10) and (8,8) CAS. The second excited state triplet provided by the singlet density lies 0.4-0.6 kcal/mol above the first excited state singlet for the (8,8)/(10,10) CAS pair. This pathway's Barrier 2 intermediate has a triplet ground state with a singlet that lies 1.7 kcal/mol above it for the (8,8)/(10,10) CAS pair.

Pathway 7 has a singlet ground state throughout. Now, the Barrier 1 intermediate is identical to the Barrier 2 in pathway 5 and thus has the lowest-lying triplet far above the

singlet ground state again. This result for this intermediate also occurs in pathways 8 and 9. The lowest-lying triplet for the transition state lies 46.4 and 44.7 kcal/mol above the ground state singlet for the (8,8) and (10,10) CAS, respectively. The Barrier 2 products in pathway 7 dissociate to doublets; this dissociation results in overall degenerate singlet and triplet surfaces.

Both pathways 8 and 9 have a singlet ground state throughout the whole path. Pathway 8 calculations do not reveal any close-lying states for the triplet of interest. As for pathway 9, it has the intermediate common to pathways 5, 7, and 8. However, the (12,12) and (14,13) CAS also give a second excited state singlet derived from the triplet density above the lowest triplet by 4.0 and 5.8 kcal/mol, respectively. Pathway 5 has the same state ordering. Barrier 2 minimum analysis is given in reaction pathway 3.

For the additional reaction pathway 10 located in this paper, intermediates have already been analyzed in earlier pathways 5 – 9. However, analysis of the transition state shows the lowest-lying triplet lies 15.0 kcal/mol above the ground state singlet. As well, a second excited singlet state derived from the triplet density again lies within a few kcal/mol of the lowest triplet state.

Conclusions:

In the current study, CASSCF and MRMP2 provide accurate barriers and minimum energy paths for the lowest-lying triplet of O + ethylene; and additional single point calculations with more states for the singlet and triplet at the triplet geometries yield preliminary surface crossing data. MR-AQCC calculations to examine additional correlation effect for reactions 1, 3, and 4 usually give differences from the MRMP2 values of less than 1 kcal/mol. Some exceptions are the difference of ~2 kcal/mol for reaction 3 Barrier 2 and

~3 kcal/mol for Barrier 1 in reaction 1. As noted earlier, the near low-lying states that are likely causing these larger differences will require other approaches, such as state averaged wavefunctions, and will be examined in future work. Figure 3 summarizes the energetics of the lowest triplet surface with the best MRMP2 values from this work. During the course of the investigation, the act of choosing an active space reveals several issues within the framework of this study. For several of the reaction pathways, root flipping shows that the converger cannot provide the most proper active space for a particular CAS size that would recover the majority of the static correlation. As a result, CASs with and without in-out correlation are constructed; and various issues and differences arise from this compromise at both the CASSCF and MRMP2 levels. However, these issues appear to be minor with respect to the energetics. As already discussed in particular, the existence of certain CASSCF stationary points gets called into question when these extrema exist only without in-out correlation; as such, the CAS constructions with in-out correlation can be considered improper for those particular reactions.

As well, MRMP2 energies on top of various CASs lead to several items of note. First, as evident from the single point MRMP2 runs, obtaining reasonable energetics for the barriers requires the inclusion of dynamic correlation in some fashion. In this study correlation from MRMP2 lowers energy barriers in many but not all cases; and in most cases, MRMP2 single point energies along the IRC tend to shift the location of energy maxima and minima and hence change the overall barrier. In particular, (12,12) and larger CASs do not experience much of a horizontal, or geometrical, shift but still experience a vertical, or energy, shift that drastically changes the barrier. A smaller CAS may or may not experience a geometrical shift with the addition of dynamic correlation. Second, the various CASs lead

to similar but not identical barriers that lie within ~1-2 kcal/mol. Third, MRMP2 corrections affect the energetics more in the cases of bimolecular products to transition state barriers than in the cases of unimolecular intermediate to transition state barriers.

As a result of the above multireference outcomes, the lowest-lying triplet barriers in this study qualitatively agree with Nguyen's barriers. However, not all barriers lie within chemical accuracy of one another. Certain barriers differ for distinct physical reasons other than the inclusion of correlation. For example, both barriers in pathway 2 and Barrier 2 in pathway 4 in this study account for hydrogen bonding; hence, a substantial difference in barriers occurs between the two studies. On the other hand, evidence suggests the dynamic correlation energetically alters Barrier 2 in pathway 1 and both barriers in pathway 6. As well, Barrier 2 in pathway 7 and Barrier 1 in pathway 8 both differ by 2.7 kcal/mol from Nguyen's corresponding barriers. In addition to these comparisons, Barrier 1 in pathway 9 provides some insight into the energetic difference between a proper active space and single, compromised active space. In the final path 10, the hydrogen shift from the far carbon to the oxygen was found.

The preliminary surface crossing analysis confirms the predominant surface crossing(s) for the system to be around the CH_2CH_2O biradical. However, several more crossings become important with hyperthermal considerations. The first and second excited states for the CH_3CHO biradical lie close to one another. As well, in the pathway 3 dissociation, a singlet state lies within ~5 kcal/mol of the lowest-lying triplet transition state. The entire reaction pathway 6 has close-lying triplet and singlet states; and the transition state itself has an additional close-lying second excited state singlet. This nearby state at least partially explains why the results from pathway 6 significantly differ from others' results.

This study also shows that some excited state characterizations require the inclusion of the O2p in the active space (e.g. pathway 2).

Future research will involve a close examination of the singlet surface, the location of the surface crossing using state averaged MCSCF, and the spin-orbit coupling near these crossings. These results will be used to fit to semi-empirical methods such as those of Granucci and Toniolo[38] and then perform dynamics.

Acknowledgments:

The authors are indebted to Michael W. Schmidt and Mark S. Gordon for help in using the capabilities of GAMESS and MCSCF. This material is based upon work supported by the National Science Foundation under Grant No. OISE-0730114 for the Partnerships in International Research and Education (PIRE) and by the Robert A. Welch Foundation under Grant No. D-0005. TeraGrid resources were provided by the Texas Advanced Computing Center (TACC). Support was also provided by the High-Performance Computing Center (HPCC) at Texas Tech University, under the direction of Philip W. Smith. In addition, this work was supported by the Austrian Science Fund within the framework of the Special Research Program F16 (Advanced Light Sources) and Project P18411-N19. TLW acknowledges computing resources purchased through funds provided by Ames Laboratory and Iowa State University.

Supporting Information Available:

Additional information including the CASSCF and relevant MR-AQCC geometries, absolute energies, frequencies, ZPE, and moments of inertia along with absolute MRMP2 energies is available free of charge via the Internet at http://pubs.acs.org.

Table 1. Description of the main CAS sizes used for each reaction pathway. Only the lone pairs and bonding orbitals are shown; the antibonding orbitals can be inferred from bonding orbitals and the indication of in-out correlation. The O2s is always in the core.

reaction pathway	active space	in-out correlation	Description
1	(6,6)	N	CCσ, CCπ, biradical(2O2p) →CCσ, COσ, biradical(farC2p,O2p)
	(8,8)	Y	CCσ, CCπ, biradical(2O2p), lone O2p →CCσ, COσ, biradical(farC2p,O2p), lone O2p
2	(12,12)	N	4 CHσ, CCπ, biradical(2O2p) →OHσ, 3 CHσ, CCπ, biradical(C2p,O2p)
3	(8,8)	N	COσ, 2 CHσ, biradical(farC2p,O2p) →COσ, CHσ, COπ, biradical(farC2p;O2p,H)
	(10,10)	Y	COσ, 2 CHσ, biradical(farC2p,O2p), lone O2p →COσ, CHσ, COπ, biradical(farC2p;O2p,H), lone O2p
4	(6,6)	N	CCσ, COσ, biradical(farC2p,O2p) →COπ, COσ, biradical(farC2p,farC2p)
	(8,8)	Y	CCσ, COσ, biradical(farC2p,O2p), lone O2p →COπ, COσ, biradical(farC2p,farC2p), lone O2p
5	(14,14)	N	4 CHσ, CCσ, COσ, biradical(farC2p,O2p) →4 CHσ, CCσ, COσ, biradical(nearC2p,O2p)
6	(8,8)	N	2 CHσ, COσ, biradical(farC2p,O2p) →COσ, CHσ, OHσ, biradical(2C2p)
	(10,10)	Y	2 CHσ, COσ, biradical(farC2p,O2p), lone O2p →COσ, CHσ, Ohσ, biradical(2C2p), lone O2p
7	(8,8)	Y	COσ, O2p, CCσ, biradical(nearC2p;O2p, O2p) →COσ, 2 COπ, biradical(nearC2p;O2p, farC2p)
8	(8,8)	Y	COσ, O2p, CHσ, biradical(nearC2p;O2p, O2p) →COσ, 2 COπ, biradical(nearC2p;O2p, H)
9	(12,12)	N	COσ, CCσ, 3 CHσ, biradical(nearC2p, O2p) →COσ, CCσ, 2 CHσ, COπ, biradical(farC2p;O2p, H)
	(14,13)	N	COσ, CCσ, 3 CHσ, biradical(nearC2p, O2p), lone O2p → N/A
10	(12,12)	N	2 CHσ, COσ, CCσ, CHσ, biradical(2C2p) →2 CHσ, COσ, CCσ, OHσ, biradical(nearC2p,O2p)

Table 2. Barriers to reactions in kcal/mol calculated at the CASSCF level without and with (in parenthesis) ZPE. Barrier 1 refers to the reactant-transition state barrier, and Barrier 2 refers to the product-transition state barrier.

reaction pathway	active space	CASSCF Barrier 1[a]	CASSCF Barrier 2[a]	Nguyen Barrier 1[b]	Nguyen Barrier 2[b]
1	(6,6)	19.2 (19.3)	17.2 (16.4)	1.3	25.3
	(8,8)	13.9 (14.3)	21.0 (20.5)		
2	(12,12)	34.4 (30.3)	13.4 (12.0)	11.0	3.5
3	(8,8)	18.7 (14.6)	14.9 (16.4)	15.3	8.3
	(10,10)	15.1 (11.2)	17.1 (18.6)		
4	(6,6)	18.1 (15.8)	17.7 (19.7)	22.1	4.7
	(8,8)	13.9 (11.8)	19.3 (21.5)		
5	(14,14)	48.3 (46.0)	51.5 (48.1)	31.5	39.9
6	(8,8)	40.3 (37.6)	42.8 (39.8)	28.8	38.5
	(10,10)	34.9 (32.4)	46.9 (43.8)		
7	(8,8)	10.3 (8.0)	22.4 (26.5)	12.8	8.3
8	(8,8)	16.3 (12.0)	20.1 (21.9)	17.4	8.5
9	(12,12)	39.8 (34.4)	32.8 (34.0)	40.4	25.0
	(14,13)	39.6 (34.1)	N/A		
10	(12,12)	44.1 (40.7)	41.6 (39.1)	N/A	N/A

a. CASSCF barriers are obtained from unconstrained optimization of intermediates and ~3-5 Å separated products.

b. Nguyen's barriers are the average barriers from different levels of theory listed in their publication.[5]

Table 3. Barriers to reactions in kcal/mol calculated at the MRMP2 level. The bare number is the SPMRMP2 barrier (see text) and always includes CASSCF ZPE, the number in parenthesis is the shifted barrier from the MRMP2 IRCs, and the number in square brackets is the shifted MRMP2 barrier including ZPE from the MCSCF maxima and minima. Barrier 1 refers to the reactant-transition state barrier, and Barrier 2 refers to the product-transition state barrier.

reaction pathway	active space	MRMP2 Barrier 1[a]	MRMP2 Barrier 2[a]	Nguyen Barrier 1[b]	Nguyen Barrier 2[b]
1	(6,6)	-1.5 (2.4) [2.5] 3.0[c]	17.1 (21.9) [21.1]	1.3	25.3
	(8,8)	0.0 (1.4) [1.8] 2.3[c]	17.5 (19.8) [19.2]		
2	(12,12)	8.9 (13.4) [9.3] 9.6[c]	4.4 (6.4) [4.9] 5.3[c]	11.0	3.5
3	(8,8)	13.6 (18.4) [14.3]	6.4 (5.6) [7.1] 7.4[c]	15.3	8.3
	(10,10)	14.0 (19.9) [15.9]	5.0 (5.5) [7.0] 7.3[c]		
4	(6,6)	17.1 (22.4) [20.2]	3.1 (4.2) [6.2] 5.8[c]	22.1	4.7
	(8,8)	16.9 (24.2) [22.1]	1.8 (4.8) [7.0] 6.8[c]		

reaction pathway	active space	MRMP2 Barrier 1[a]	MRMP2 Barrier 2[a]	Nguyen Barrier 1[b]	Nguyen Barrier 2[b]
5	(14,14)	33.1 (35.4) [33.1]	40.0 (43.4) [40.0]	31.5	39.9
6	(8,8)	25.7 (28.4) [25.7]	34.1 (37.2) [34.1]	28.8	38.5
	(10,10)	26.7 (29.4) [26.8]	35.6 (38.9) [35.7]		
7	(8,8)	8.9 (15.2) [12.9]	6.5 (6.4) [10.6] 11.0[c]	12.8	8.3
8	(8,8)	16.4 (24.4) [20.1]	5.7 (7.6) [9.4] 9.6[c]	17.4	8.5
9	(12,12)	39.1 (45.0) [39.6]	24.7 (24.0) [25.2]	40.4	25.0
	(14,13)	37.1 (42.6) [37.1]	N/A		
10	(12,12)	27.6 (31.0) [27.6]	28.4 (31.0) [28.4]	N/A	N/A

a. Single point MRMP2 (SPMRMP2) values derive from CASSCF barriers, which are obtained from unconstrained optimization of intermediates and ~3-5 Å separated products.

b. Nguyen's barriers are the average barriers from several different levels of theory that include ZPE listed in their publication.[5]

c. MRMP2 barrier with ZPE from shifts in minima in addition to shifts in transition states.

Table 4. MR-AQCC and shifted MRMP2 barriers (with shifted transition state only) in kcal/mol given for selected reactions. The MR-AQCC results are shown with BS1, BS2 and BS3 from top to bottom in each cell of the table. All barriers are reported without ZPE.

reaction pathway	active space	MRMP2 Barrier 1	MRMP2 Barrier 2	MR-AQCC Barrier 1	MR-AQCC Barrier 2
1	(6,6)	(2.4)	(21.9)	4.2 3.5 1.1	19.4 22.1 22.6
	(8,8)	(1.4)	(19.8)		
3	(8,8)	(18.4)	(5.6)	18.4 17.9 20.6	8.8 8.9 7.4
	(10,10)	(19.9)	(5.5)		
4	(6,6)	(22.4)	(4.2)	22.4 24.2 25.2	6.7 6.2 4.9
	(8,8)	(24.2)	(4.8)		

Figure 1. An example of an unphysical, in-out correlation, antibonding orbital found in active spaces that contain three O2p orbitals.

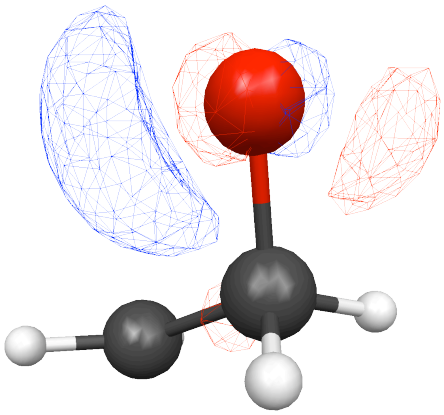


Figure 2. (8,8) MCSCF and MRMP2 IRCs for Pathway 1 display the shift in barriers. The plot below has the highest energies along the IRC for both levels of theory set equal.

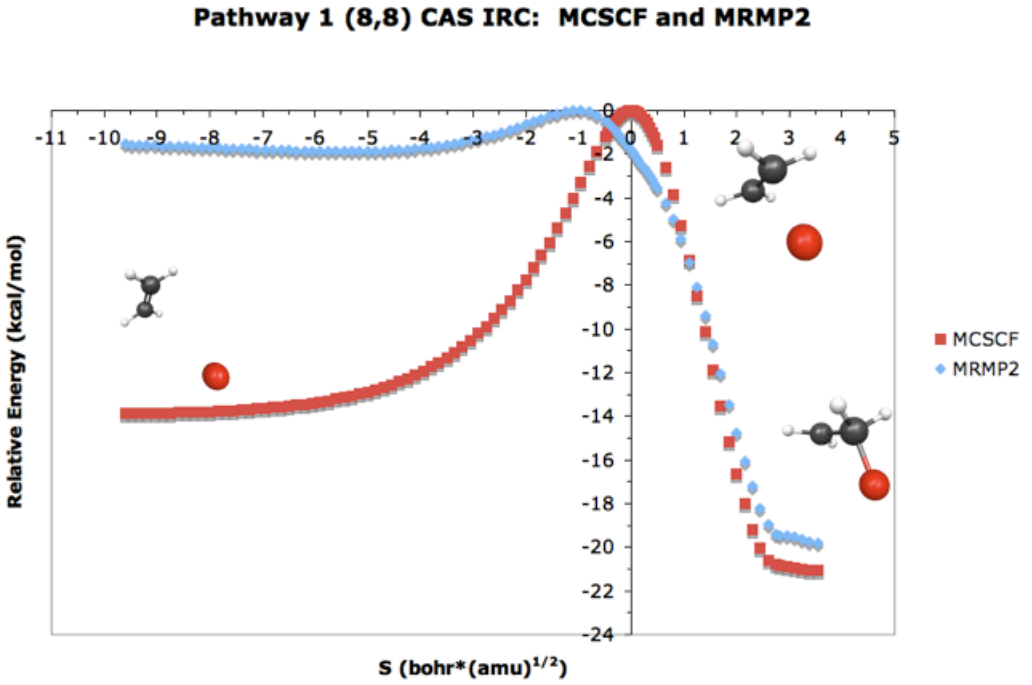
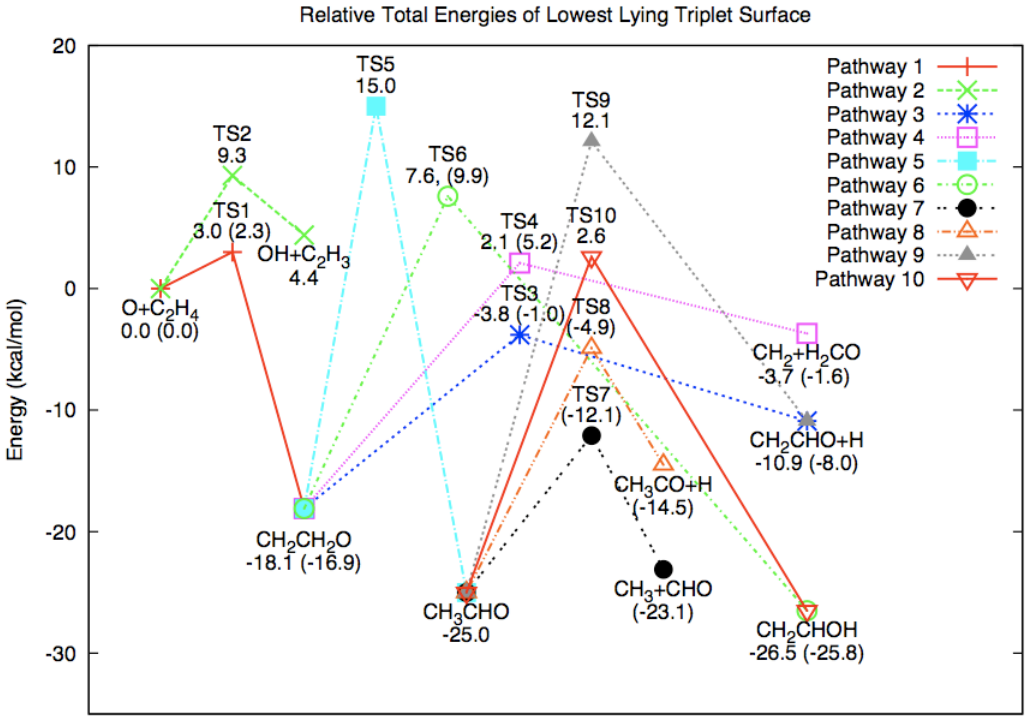


Figure 3. Total MRMP2 energies relative to reactants O + C_2H_4 for the overall lowest-lying triplet surface. All relative energies are constructed in a piece-wise fashion to give an energy relative to the initial reactants. Values in parentheses derive from a CAS with in-out correlation while plain values derive from a CAS with no in-out correlation.



References:

[1]Gardiner, W. C., Jr.; Editor *Combustion Chemistry*; Springer-Verlag, New York, **1984**.

[2]Casavecchia, P.; Capozza, G.; Segoloni, E.; Leonori, F.; Balucani, N.; Volpi Gian, G. *J Phys Chem A* **2005**, *109*, 3527.

[3]Su, H.; Zhao, S.; Liu, K.; Xiang, T. *J Phys Chem A* **2007**, *111*, 9600.

[4]Lee, S.-H.; Huang, W.-J.; Chen, W.-K. *Chem. Phys. Lett.* **2007**, *446*, 276.

[5]Nguyen, T. L.; Vereecken, L.; Hou, X. J.; Nguyen, M. T.; Peeters, J. *J. Phys. Chem. A* **2005**, *109*, 7489.

[6]Yamaguchi, K.; Yabushita, S.; Fueno, T.; Kato, S.; Morokuma, K. *Chem. Phys. Lett.* **1980**, *70*, 27.

[7]Dupuis, M.; Wendoloski, J. J.; Takada, T.; Lester, W. A., Jr. *J. Chem. Phys.* **1982**, *76*, 481.

[8]Fueno, T.; Takahara, Y.; Yamaguchi, K. *Chem. Phys. Lett.* **1990**, *167*, 291.

[9]Jursic, B. S. *THEOCHEM* **1999**, *492*, 85.

[10]Smith, B. J.; Nguyen Minh, T.; Bouma, W. J.; Radom, L. *J. Am. Chem. Soc.* **1991**, *113*, 6452.

[11]Hu, W.; Lendvay, G.; Maiti, B.; Schatz, G. C. *J. Phys. Chem. A* **2008**, *112*, 2093.

[12]Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

[13]Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822.

[14]Mebel, A. M.; Morokuma, K.; Lin, M. C. *J. Chem. Phys.* **1995**, *103*, 7414.

[15]Werner, H. J.; Knowles, P. J. *J. Chem. Phys.* **1988**, *89*, 5803.

[16]Knowles, P. J.; Werner, H. J. *Chem. Phys. Lett.* **1988**, *145*, 514.

[17]Gordon, M. S.; Schmidt, M. W. *Theory Appl. Comput. Chem.: First Forty Years* **2005**, 1167.

[18]Roos, B. *Adv. Chem. Phys.* **1987**, *69*, 399.

[19]Ruedenberg, K.; Sundberg, K. R. *Quantum Sci.* **1976**, 505.

[20]Schmidt, M. W.; Gordon, M. S. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233.

[21]Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

[22]Ivanic, J.; Ruedenberg, K. *Theor. Chem. Acc.* **2001**, *106*, 339.

[23]Lengsfield, B. H., III. *J. Chem. Phys.* **1980**, *73*, 382.

[24]Fletcher, G. D. *Mol. Phys.* **2007**, *105*, 2971.

[25]Yarkony, D. R. *Chem. Phys. Lett.* **1981**, *77*, 634.

[26]Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1980**, *72*, 880.

[27]Boys, S. F. *Quantum Theory At., Mol., Solid State.* **1966**, 253.

[28]Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154.

[29]Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.

[30]Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.

[31]Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481.

[32]Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Muller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664.

[33]Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Mueller, T.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Boehm, H. J.; Chang, A.; Comeau, D. C.; Gdanitz, R.; Dachsel, H.; Ehrhardt, C.; Ernzerhof, M.; Hoechtl, P.; Irle, S.; Kedziora, G.; Kovar, T.; Parasuk, V.; Pepper, M. J. M.; Scharf, P.; Schiffer, H.; Schindler, M.; Schueler, M.; Seth, M.; Stahlberg, E. A.; Zhao, J.-G.; Yabushita, S.; Zhang, Z.; Barbatti, M.; Matsika, S.; Schuurmann, M.;

Yarkony, D. R.; Brozell, S. R.; Beck, E. V.; Blaudeau, J.-P. **2006**, www.univie.ac.at/columbus.

[34]Shepard, R.; Shavitt, I.; Pitzer, R. M.; Comeau, D. C.; Pepper, M.; Lischka, H.; Szalay, P. G.; Ahlrichs, R.; Brown, F. B.; Zhao, J. G. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1988**, *22*, 149.

[35]Shepard, R.; Lischka, H.; Szalay, P. G.; Kovar, T.; Ernzerhof, M. *J. Chem. Phys.* **1992**, *96*, 2085.

[36]Fonderie, V.; Maes, D.; Peeters, J. "The kinetic coefficient of the ethylene + atomic oxygen reaction over extended pressure and temperature ranges," Dep. Chem.,Kathol. Univ. Leuven,Louvain,Belg., **1984**.

[37]Baulch, D. L.; Cobos, C. J.; Cox, R. A.; Frank, P.; Hayman, G.; Just, T.; Kerr, J. A.; Murrells, T.; Pilling, M. J.; et al. *J. Phys. Chem. Ref. Data* **1994**, *23*, 847.

[38]Granucci, G.; Toniolo, A. *Chem. Phys. Lett.* **2000**, *325*, 79.

CHAPTER 3. O + C_2H_4 POTENTIAL ENERGY SURFACE: EXCITED STATES AND BIRADICALS AT THE MULTIREFERENCE LEVEL

[With kind permission from Springer Science+Business Media: *Theor. Chem. Acc.,* O + C_2H_4 potential energy surface: excited states and biradicals at the multireference level, 131, 2012, 1123, West, A. C.; Lynch, J. D.; Sellner, B.; Lischka, H.; Hase, W. L.; Windus, T. L.; 1-14, Copyright Springer-Verlag 2012]

Aaron C. West, Joseph D. Lynch, Bernhard Sellner, Hans Lischka, William L. Hase, and Theresa L. Windus

Abstract:

The focus of this study is to understand the multiconfigurational nature of the biradical species involved in the early reaction paths of the oxygen plus ethylene PES. In previous work (*J. Phys. Chem. A* **2009**, *113*, 12663)[1], the lowest-lying O(^{3}P) + C_2H_4 PES was extensively explored at the MCSCF, MRMP2, and MR-AQCC levels of theory. In the current work, ground and excited, triplet and singlet state reaction paths for the initial addition of oxygen to ethylene were found at the MCSCF and MRMP2 levels along with five singlet pathways near the •CH_2CH_2O• biradical at the MCSCF, MRMP2, and CR-CC(2,3) levels. One of these five paths can lead to the CH_2CO + H_2 products from CH_3CHO rather than from the •CH_2CH_2O• biradical, and this pathway was investigated with a variety of CAS sizes. To provide further comparison between the MRMP2 and CR-CC(2,3) levels, MR-AQCC single point energies and optimizations were performed for select geometries.

After the initial exploration of this region of the surface, the lowest singlet-triplet surface crossings were explicitly determined at the MCSCF level. Additional MRMP2 calculations were performed to demonstrate the limitations of single state perturbation theory in this biradical region of the PES, and SO-MCQDPT2 single point energies using SA-MCSCF were calculated on a grid of geometries around the primary surface crossing. In particular, these calculations were examined to determine a proper active space and a physically reasonable number of electronic states. The results of this examination show that at least four states must be considered to represent this very complex region of the PES.

Introduction:

Understanding the kinetics and dynamics of the O + ethylene reaction is important for both hydrocarbon combustion studies[2] and development of materials that resist surface erosion for space vehicles in low-Earth orbit. Several low pressure crossed molecular beam experiments[3-5] were performed for this reaction. On the lowest-lying triplet potential energy surface (PES), $O(^3P)$ addition to ethylene first forms a ketocarbene that connects to reaction pathways leading to the three most dominant set of products $CH_2CHO + H$, $CH_3 + CHO$, and $CH_2 + H_2CO$[6]. The pathways to the CH_2CHO and H_2CO products are low barrier, single-step decompositions, and these reactions occur under thermal conditions. In contrast, the CHO reaction channel has a large rearrangement barrier prior to decomposition and is only observed for hyperthermal reactant conditions. In the vicinity of the triplet ketocarbene, triplet-singlet non-adiabatic electronic transitions are possible to the singlet PES. If this electronic transition occurs, a hydrogen shift leads to the $CH_3 + HCO$ products.

The overall product distributions are nearly pressure independent over a temperature range of 287-2000 K and a pressure range of 0.007-1 atm[6]. Nguyen et al.[6] have shown that

collisional stabilization of the ketocarbene adduct remains negligible (<10%) even at 100 atm and ~1500 K. Thus, the product distributions are primarily controlled by the temperature. As found from both theoretical calculations[6] and experiment[3], at 298 K the percentage yields for the three dominant channels (i.e. CH_2CHO + H, CH_3 + CHO, and CH_2 + H_2CO) are 40±10, 50±10, and 10±5, respectively. For the much higher temperature of 2000 K, theoretical calculations give percentages of 18.5, 36.9, and 29.1 for the respective yields of these three products[6]. As the temperature is varied, both the product yields and distributions change.

Several initial theoretical studies of this reaction are present in the literature[7-11]. This work was followed by the calculations of Nguyen et al.[6] considered above and those of Hu et al.[12], who studied the reaction in substantial detail at different levels of theory on both the triplet and singlet PESs. In addition, Joshi et al.[13] studied oxirane decomposition with B3LYP/6-311++G(d,p) optimized geometries and G3B3/6-311++G(d,p) single point energies and investigated the biradical region of the singlet PES with an unrestricted singlet wavefunction and spin-projection. Yang et al.[14] located ~100 stationary points involved in reactions that emanate from the acetaldehyde potential energy minimum. They used B3LYP/6-311++G(d,p) geometries and CCSD(T)/cc-pVTZ single point energies with basis set superposition error corrections for weakly bound complexes. Recent studies at the CCSD(T)/aug-cc-pVTZ level of theory by Bowman investigated how pathways important to the CH_4 + CO product channel bypass the minimum energy path (MEP) and hence the traditional transition state (TS)[15-17]. Combining both experimental and theoretical results, another recent study[18] proposed a novel pathway (i.e. $\bullet CH_2CH_2O\bullet \rightarrow CH_2CO + H_2$). However, the theoretical component is somewhat uncertain since it consisted of G3(MP2)

single state, single reference results with large spin contamination. To date, most theoretical studies on this system rely on single determinant methods since understanding even the most important reactions for this system proves to be a challenge for multireference electronic structure theory methods. The calculations presented here examine the multireference nature of the O + C_2H_4 reactions, which show several important non-adiabatic transitions in the biradical region.

For the O + ethylene reaction, both the two lowest-lying triplets and two lowest-lying singlets play important roles in the reactions of the initially formed biradicals and the ensuing dynamics. Our initial study[1] of the lowest-lying triplet PES examined pathways 1 – 9 from the Nguyen et al.[6] work as well as an additional path 10. The current study expands this earlier work[1] by examining multireference barriers for the lowest-lying singlet paths 10 – 13, 20 as labeled in Nguyen et al.[6], excited states for path 1, the Miyoshi path (M), and important surface crossings. In particular, this study examines the following reactions:

Triplet A′ 1) $\bullet\bullet O(^3P) + C_2H_4 \rightarrow \bullet CH_2CH_2O\bullet$

Singlet A_1 1) $\bullet\bullet O(^1D) + C_2H_4 \rightarrow$ oxirane

Singlet A″ 1) $\bullet\bullet O(^1D) + C_2H_4 \rightarrow \bullet CH_2CH_2O\bullet$

10) $\bullet CH_2CH_2O\bullet \rightarrow CH_3CHO$

11) $\bullet CH_2CH_2O\bullet \rightarrow$ oxirane

12) $\bullet CH_2CH_2O\bullet \rightarrow CH_2CHOH$

13) $CH_3CHO \rightarrow$ oxirane

M) $\bullet CH_2CH_2O\bullet \rightarrow CH_2CO + H_2$

20) $CH_3CHO \rightarrow CH_2CO + H_2$

Methods:

Results from calculations at the multiconfigurational self-consistent field (MCSCF) level in combination with several dynamical correlation methods were obtained with the GAMESS software[19]. Since the multireference approach for these calculations has already been explained for the triplet surface[1], only a summary of that procedure is given here. To recover the majority of the quasi-degenerate correlation, all calculations were carried out at the full optimized reaction space (FORS)[20-22], or complete active space self-consistent field (CASSCF), level of theory using the aug-cc-pVTZ[23] basis set. The previous paper[1] compared calculations between aug-cc-pVTZ and aug-cc-pVDZ[23] and showed the triple zeta basis set to be sufficient for this system. The determinant-based method[24] and the full Newton-Raphson converger with the augmented Hessian technique[25-27] were used because of the large complete active space (CAS) sizes and the biradical species. Restricted Hartree-Fock (RHF) or restricted open shell Hartree-Fock (ROHF) calculations with modified valence orbitals[28] and Boys localization[29] provided good starting orbitals in most situations. CASSCF stationary point searches for species in the reactions described above employed analytic gradients and central differenced, numerical Hessians. As in the first paper[1], second order, Gonzalez-Schlegel (GS2)[30], intrinsic reaction coordinate (IRC) calculations identified minima associated with each TS.

Once a reaction path was located, single state multireference Møller-Plesset second order perturbation theory (MRMP2)[31] as well as completely renormalized coupled-cluster singles, doubles, and non-iterative triples (CR-CC(2,3))[32,33] single point energies using the aug-cc-pVTZ basis set were performed along the IRC to recover the majority of the dynamic correlation. In the previous work on the triplet surface, MRMP2 results quantitatively agreed

with multireference average quadratic coupled cluster (MR-AQCC)[34] results obtained using COLUMBUS[35-37]. Since both analytic gradients and Hessians at the MRMP2 level are unavailable in GAMESS, numerical derivatives are required. Unfortunately, numerical derivatives involve computation at asymmetric geometries where the multireference wavefunction might converge to any close-lying states by orbital root flipping (ORF) to a different active space. Thus, given that the MRMP2 results depend on the active space, geometry optimizations, IRCs, and Hessians at this level are numerically prohibitive and sometimes unfeasible.

To further validate the results, the COLUMBUS program[35-37] was used for GVB-PP and MR-AQCC single point energy and geometry optimization calculations at select geometries with the cc-pVDZ[23] basis. The GVB-PP(9) approach utilizes the generalized valence bond-perfect pairing (GVB-PP) method with 9 electron pairs[38-40]. Besides the chemical core O1s and C1s orbitals, which are doubly occupied, all other orbitals are treated on an equal footing allowing for the correct dissociation of individual electron pairs. The total electronic wave function consists of an antisymmetrized product of the nine perfect pairs (PPs), yielding a space of 2^9 (512) configurations. For the MR-AQCC calculations, a PP(5) reference space was chosen for path 12 and a PP(4) reference space for path 13. The choice of pairs was guided by considerations of the necessary orbitals for the reaction, balanced treatment of bonds, and computational cost. The bonding orbitals of the remaining pairs are doubly occupied in the reference space. For path 12, these doubly occupied orbitals represent two CH, the CC, and the CO bond while in path 13 the four CH and the CC bonds are used. Based on these reference spaces, single and double excitations are allowed by imposing generalized interacting space restrictions[41]. With 32 (16) reference configuration

state functions for PP(5) (PP(4)), the total MR-AQCC expansion space consists of 3,088,992 (1,561,728) configurations.

Even though most of the lowest-lying singlet PES lies far below the lowest-lying triplet surface, the $\bullet CH_2CH_2O\bullet$ biradical has several degenerate electronic states that can lead to non-radiative transitions and to several major singlet minima. To understand these electronic states, two minimum energy crossing (MEX) regions were located at the CASSCF level with GAMESS: one near the $\bullet CH_2CH_2O\bullet$ triplet minimum and one near the $\bullet CH_2\bullet CHOH$ triplet biradical. Because calculations based on a single state (8,7) CAS resulted in ORF, the crossing seam near $\bullet CH_2CH_2O\bullet$ was located with a (6,6) CAS, which yielded equal energy gaps at the CASSCF and MRMP2 levels between the lowest-lying triplet and lowest-lying singlet states at the triplet $\bullet CH_2CH_2O\bullet$ minimum; this specification ensures that both surfaces receive similar amounts of correlation at these levels of theory.

Investigating these PES regions requires both location of the crossing seams[42-44] and a multiple state perturbation approach[45-48] with spin orbit coupling (SOC)[49,50]. Toward this end, second order, spin orbit multiconfigurational quasi-degenerate perturbation theory (SO-MCQDPT2)[50,51] single point energies were calculated with GAMESS at the CASSCF MEX geometries. The SO-MCQDPT2 level of theory combines spin orbit effects of the partial or full Breit-Pauli operator[51-54] with MCQDPT2[48]. The SO calculations included full 1 electron and partial 2 electron integrals, which result in a mean field adjustment to the 1 electron integrals. SO-CAS configuration interaction (SO-CASCI) energies are also reported since they are included in the SO-MCQDPT2 energy evaluation. The second order, perturbative calculations included all excitations from the reference into the virtual space from all but the chemical core orbitals (O1s, C1s). Unlike the calculations with pure (i.e. single) state

solutions, state averaged (SA)-MCSCF can lead to appropriate active spaces that include the lone O2p without in-out correlation (IOC). So, with the proper electronic states, the SO-MCQDPT2 approach can lead to physically meaningful results.

To further investigate the multi-state nature of the $\bullet CH_2CH_2O\bullet$ biradical region, (8,7) SA-CASSCF and SO-MCQDPT2, single point energies were computed for the geometries previously obtained in the CASSCF IRCs near the biradical region. The number of electronic states included was primarily determined by examination of the energy gaps between the states for both pure and SA wavefunctions. In addition, around the $\bullet CH_2CH_2O\bullet$ MEX geometry, a series of (8,7) SA CASSCF and corresponding SO-MCQDPT2, single point energies using geometries obtained from (6,6) CAS, triplet constrained optimizations were performed to further describe this very flat, 14 dimensional intersection seam.

Results and Discussion:

Table 1 displays the active space for each reaction where the label 'far' refers to the carbon farthest from oxygen. Examination of bond length changes, natural orbital occupation numbers, and localized molecular orbital coefficients in different CAS sizes over the course of each reaction ultimately decide the choice of the final active space. Tables 2 and 3 give the overall barrier information for all elementary reactions listed in the Introduction where each TS divides each pathway into a forward barrier (reactant to TS) and a reverse barrier (product to TS). Table 2 shows the CAS barriers with and without zero point energies (ZPE). Table 3 contains energetics at several levels of theory: 1) the barriers from single point MRMP2 energies along the CASSCF IRCs; 2) optimized MR-AQCC barriers for paths 12 and 13; 3) Nugyen et al.[6] and Yang et al.[14] barriers; 4) single point MR-AQCC barriers along the CASSCF IRC for paths 12 and 13; and 5) the barriers from single

point CR-CC(2,3) energies along the CASSCF IRCs. While the single point energies along the CASSCF IRC geometries are not strictly barriers, they should approximate the barriers[1]. Differences between single point MRMP2 data at the MCSCF minima and maxima are labeled stationary point MRMP2 (SPMRMP2) energy barriers and always include ZPE (except for species with ORF issues in numerical Hessians as discussed below). However, as shown below, single point MRMP2 energies along the IRC can result in shifted barriers because of shifts in the TS and/or the minimum(a) (i.e. reaction coordinate/horizontal shifts and energetic shifts).

As in our first paper[1], results with IOC tend to incorporate more dynamic correlation into the reference wavefunction. However, if a weakly occupied IOC orbital undergoes ORF with a virtual orbital and if dynamic correlation has a large contribution to the total correlation, then barriers with different IOC orbitals do not necessarily lead to accurate energetics. Therefore, IOC orbitals in the active space might result in configurations that lead to unphysical barriers. As a result, unless either the IOC constitutes proper radial correlation at some geometry in the pathway or ORF prevents placing the lone O2p in the MCSCF core, CASSCF active spaces do not contain IOC orbitals.

For the elementary reaction results in this section, the descriptions contain the following, unless otherwise explicitly noted: comparisons of barriers and energetics include MCSCF ZPE; the CASSCF active spaces do not contain IOC and contain one lone O2p orbital in the MCSCF core; any reported MRMP2 barriers include all occurring reaction coordinate shifts; any reported CR-CC(2,3) barriers in the text do not include any shifts, are determined at the TS and minima from the MCSCF IRC, and are listed with and without MCSCF ZPE; MR-AQCC barriers in the text always include MR-AQCC ZPE; any

references made to stationary points refer to CASSCF stationary points (except for MR-AQCC optimizations); and individual pathways report only non-negligible geometry changes in CASSCF stationary points relative to Nguyen's[6] B3LYP and/or CASSCF results.

Triplet A′ Pathway 1: $\bullet\bullet O + C_2H_4 \rightarrow \bullet CH_2CH_2O\bullet$

This reaction has an A′ state in C_s symmetry throughout most of the pathway and is well represented by a (6, 6) CAS. The reaction begins with an initial oxygen approach on the first excited state and then passes the A′ symmetric saddle point into the biradical region. Like the triplet A″ ground state pathway[1], this excited state path becomes asymmetric very close to the biradical minimum. The current methodologies of MCSCF usually do not allow for the determination of pure, asymmetric excited states. So, in order to obtain the IRC in this region, the calculations continued to use A′ symmetry and led to a symmetric stationary point that is not a minimum and contains a single, imaginary frequency of -401 cm^{-1} that indicates that the OCCH dihedral angle bends out of the C_s plane. Also, including dynamic correlation leads to large changes in energetics and large shifts along the reaction coordinate. As a result, the forward barrier is 19.6 (3.3) kcal/mol at the CASSCF (MRMP2) level. This CAS barrier versus the MRMP2 barrier demonstrates the importance of dynamic correlation for this path. Other pathway discussions will not contain any further, unnecessary comparisons of CAS and MRMP2 barriers (see Tables 2 and 3).

Since the calculations could not fully determine the reverse barrier because of the symmetry-constrained IRC, the symmetric product provides an approximation. Given that the A″ geometries parallel the A′ geometries and the ground state A″ CASSCF ZPE changes the ground state reverse barrier by less than 1.0 kcal/mol[1], the excited state reverse barrier of

21.2 kcal/mol (without ZPE) at the MRMP2 level at the symmetric CASSCF stationary point nicely agrees with Nguyen's reverse barrier of 20.9 kcal/mol (with ZPE).

Singlet A_1 Pathway 1: $\bullet\bullet O + C_2H_4 \rightarrow$ oxirane

The calculations for this pathway used a (6,6) CAS with an A_1 state in C_{2v} symmetry throughout the IRC. Because numerical Hessian steps converge to different states in this case, Hessian calculations with a double zeta basis (which allows for analytic Hessian calculations in GAMESS) provided an approximate curvature for a triple zeta, saddle point search. Single state MRMP2 leads to barrierless, oxygen addition for this excited state with an overall reaction exothermicity of -121.5 kcal/mol.

Singlet A″ Pathway 1: $\bullet\bullet O + C_2H_4 \rightarrow \bullet CH_2CH_2O\bullet$

A (6,6) CAS calculation was also used to study this pathway, which has an A″ state in C_s symmetry only throughout the forward barrier. The dihedral rotation breaks the symmetry near the $\bullet CH_2CH_2O\bullet$ biradical as discussed in the previous section. At the (6,6) CASSCF level of theory, the two lowest, singlet states appear to intersect one another near the forward barrier A″ saddle point, and again, problems with the numerical Hessian required a double zeta Hessian approximation. Including dynamic correlation results in a barrierless, overall energy change of -61.7 kcal/mol (without ZPE).

Pathway 10: $\bullet CH_2CH_2O\bullet \rightarrow CH_3CHO$

This pathway involves a H-atom shift and requires the use of a (14,14) CAS. It is one of several pathways with a TS (see Figure 1) near the $\bullet CH_2CH_2O\bullet$ biradical and has several dominant determinants contributing to the wavefunction. The majority of path 10 is asymmetric with only acetaldehyde having an A′ state in C_s symmetry. While including some dynamic correlation does not lead to any substantial energy shifts along the reaction

coordinate as shown in Figure 2, it causes the forward barrier to change from a shallow height of 1.0 kcal/mol at the CASSCF level to a barrierless process. Therefore, no reverse barrier exists at the MRMP2 level along these CASSCF IRC points. The resultant barrierless reaction has an exothermicity of -77.6, -89.9, and -93.3 kcal/mol at the MRMP2, CR-CC(2,3), and CR-CC(2,3) without MSCF ZPE levels of theory, respectively. This energy difference shows the importance of including more dynamic correlation such as with CR-CC(2,3). It also indicates that other pathways (for example, singlet A″ pathway 1) can contribute to path 10 products in a barrierless fashion if the $\bullet CH_2CH_2O\bullet$ biradical is in that particular pathway. This conclusion applies to any other pathways whereby dynamic correlation removes the reaction barrier.

During the hydrogen shift, the CCH angle of the action CH distance changes from 110 ^{o} (CASSCF biradical) to 92 ^{o} (CASSCF TS) in the forward barrier while the CO distance changes from 1.34 to 1.22 Å in the reverse barrier. Here and throughout this paper, "action" coordinates refer to internal coordinates that significantly change in length or degree relative to remaining (i.e. "non-action") internal coordinate changes, which are defined by cutoffs. Nguyen et al. [6] report geometries at the CASSCF(8,8)/cc-pVDZ level. All geometry differences between their study and this one remain negligible.

Pathway 11: $\bullet CH_2CH_2O\bullet \rightarrow$ oxirane

This pathway gives the ring closure of the $\bullet CH_2CH_2O\bullet$ biradical to oxirane and is well described by a (6,6) CAS. The CC distance changes by only ~0.03 Å across the entire reaction pathway in the (6,6) CAS, which includes the CC σ; thus, a (4,4) CAS could represent this reaction. However, since the initial part of this pathway requires SOC in the biradical region, a (6,6) CAS is appropriate and not much more expensive than a (4,4) CAS.

This reaction has C_{2v} symmetry only for the oxirane minimum. Including dynamic correlation leads to results similar to the results for pathway 10. The reaction is essentially barrierless with a forward barrier of 0.2, -0.1, and -0.3 kcal/mol whereas the reverse reaction has a barrier of 56.6, 64.5, and 68.6 kcal/mol at the MRMP2, CR-CC(2,3), and CR-CC(2,3) without MCSCF ZPE levels, respectively.

With several determinants contributing to the wavefunction, this TS also has biradical character and geometrically resembles the starting biradical, $\bullet CH_2CH_2O\bullet$. Both the biradical minimum and TS geometries have a CO (CC) distance of 1.41 (1.51) Å. The OCC angle is 116 ^{o} in the minimum structure and 115 ^{o} in the TS, and the smaller OCCH dihedrals for both structures are 14 ± 1 ^{o}. However, the OCCH dihedral for the minimum is 176 ^{o} and for the TS is 148 ^{o}. Oxirane has a CO distance of 1.44 Å, a CC distance of 1.48 Å, and a COC angle of 62 ^{o}. In this study both dihedral angles in the TS geometry significantly differ from Nguyen et al.[6] B3LYP (CASSCF) values of 37 ^{o} (26 ^{o}) and 130 ^{o} (129 ^{o}) for the TS. As will be shown in the surface crossings section, this (6,6) CAS consistently recovers correlation in the biradical region as opposed to an (8,8) CAS with IOC. This factor strongly affects the accuracy of the forward barrier. Indeed, these structural differences emphasize the need for careful determination of the proper CAS sizes for this very complex region close to the $\bullet CH_2CH_2O\bullet$ biradical.

Pathway 12: $\bullet CH_2CH_2O\bullet \rightarrow CH_2CHOH$

With a (10,10) CAS, this pathway is dominated by a hydrogen shift. The majority of this pathway does not have symmetry; vinyl alcohol has C_s symmetry. At the MRMP2 level, this reaction is almost barrierless with a forward barrier of 0.4 kcal/mol. Hence, this pathway represents the third reaction to originate from $\bullet CH_2CH_2O\bullet$ and to become barrierless with

application of MRMP2, single point energies. Furthermore, the shifted, CR-CC(2,3) energies for the forward barrier are also small at 0.7 kcal/mol. However, unlike pathways 10 and 11, the MRMP2 and CR-CC(2,3) single point energies along the CASSCF IRC for pathway 12 still has some curvature that resembles the CASSCF IRC near the TS instead of appearing totally barrierless (see path 10, 11 and 12 IRC plots in the Supporting Information). The reverse barrier is 64.1, 71.4, and 75.9 kcal/mol at the MRMP2, CR-CC(2,3), and CR-CC(2,3) without MCSCF ZPE levels, respectively. At the MR-AQCC level, the forward barrier is 1.8 kcal/mol while the reverse barrier is 69.7 kcal/mol. However, for this theory level, the TS displays a first imaginary frequency of -1241 cm^{-1} and a second imaginary frequency of -91 cm^{-1} that could not be eliminated.

Given both discrepancies among the various levels of theory and computational limitations, single point energies using first and second order CI (CISD) with Davidson +Q correction[55,56] from a (4,4) CAS (MR-CI+Q) and aug-cc-pVDZ basis, both with and without additional excitations out of the core non-1s orbitals, were performed on the forward barrier via the occupation restricted multiple active space (ORMAS) [57] implementation in GAMESS. At the same geometries, CR-CC(2,3) calculations were also performed with the aug-cc-pVDZ basis for comparison purposes. Without ZPE, these single point energy results give forward barriers of 1.7, 9.8, -1.0, and 9.8 kcal/mol for MR-CI, MR-CI without the additional excitations, MR-CI+Q, and MR-CI+Q without the additional excitations, respectively, and -4.4 (-6.3) kcal/mol for CR-CC(2,3) with the aug-cc-pVDZ (aug-cc-pVTZ) basis. Combined with additional single point energies at these geometries (Table 3) and the above discussion for path 12, these results strongly support that the forward barrier is in fact barrierless.

The biradical character of the TS causes it to mostly resemble the initial, $\bullet CH_2CH_2O\bullet$ biradical. As the reaction progresses from the biradical to the TS, the action HCO angle changes from 104 ^{o} to 74 ^{o}, and the non-action HCCH dihedral changes from 60 ^{o} to 21 ^{o}. The CC distance changes from 1.51 to 1.35 Å during the course of the full reaction.

Pathway 13: $CH_3CHO \rightarrow$ oxirane

This portion of the PES is very difficult to characterize given that the concerted hydrogen shift with CO bond formation most likely involves several asymmetric, quasi-degenerate states. In more general terms, this pathway requires not only a transformation among lone pair and bond-antibond orbitals but also a change in dominant configurations over a small range of nuclear geometries. Because of the extremely flat, MEP curvature around the biradical-like TS, locating this stationary point proved difficult in calculations both with and without analytic Hessians. Several CAS sizes show one imaginary frequency after geometry optimization. However, depending on the initial geometry, active spaces without either the O2p or the CC σ, σ* orbitals can lead to a TS with a reasonable, distinct, single imaginary frequency that does not lead to acetaldehyde via a traditional IRC. Instead, the IRC calculations led to path 11 stationary points (sometimes with convergence issues far into the IRC calculation). Hence, even though such configuration expansions result in the proper curvature for the TS, they do not necessarily give the desired IRC (e.g. pathway 13).

This path must have at least a (16,16) CAS in order to include all quasi-degenerate effects. Since a (16,16) CAS is very computationally expensive, a (10,10) CAS with the non-action CH orbitals in the core provides a compromised solution. Because this path is so difficult to follow, GVB-PP calculations provided an initial TS geometry with the CH_2 torsion dominating the imaginary frequency. The resulting (10,10) CAS IRC gives the

correct reaction endpoints. Since the (10,10) CAS is a compromise, another (12,12) CAS was examined to provide a comparison for the energetics. The MCSCF core of the (12,12) CAS contains the O2s, CC σ, and the lone O2p orbitals in addition to the chemical core orbitals and produces stationary point energetics comparable to those obtained with the (10,10) CAS. However, this (12,12) CAS gives incorrect IRC endpoints. At the SPMRMP2 level, the forward (reverse) barrier is 78.3 (52.5) kcal/mol for the (10,10) CAS and 79.3 (51.1) kcal/mol for the (12,12) CAS. The CR-CC(2,3) (CR-CC(2,3) without MCSCF ZPE) calculations resulted in a forward barrier of 91.2 (94.2) kcal/mol and a reverse barrier of 63.8 (68.1) kcal/mol. Based on the MRMP2 and CR-CC(2,3) results, dynamic correlation changes the barrier heights but does not result in any horizontal shift in the location of the TS. At the MR-AQCC level, the forward barrier is 83.1 kcal/mol while the reverse barrier is 51.7 kcal/mol. Also, at this theory level, a Hessian calculation on TS 13 produced a -535 cm^{-1} imaginary frequency with the same qualitative nature as the corresponding GVB-PP and CAS imaginary frequencies.

TS 13 lies along a particularly flat region of the potential; this characteristic flatness distinctly shows up in the IRC. In addition, MRMP2 (MR-AQCC) energies with the aug-cc-pVTZ (cc-pVDZ) basis give an overall reaction enthalpy of 25.8 (31.2) kcal/mol, which compares well to an experimental enthalpy of 27.1 kcal/mol [58-60].

Multiple geometrical changes occur for this reaction. In particular, this TS resembles the TSs for paths 10 – 12 and also has several dominant configurations contributing to the wavefunction. For the TS, only slight differences between this study and Nguyen's study exist for the two HCC angles for the hydrogens closest to the oxygen.

Pathway Miyoshi M and 20: $\bullet CH_2CH_2O\bullet \rightarrow CH_2CO + H_2$ and $CH_3CHO \rightarrow CH_2CO + H_2$

As shown above, representations of the pathways through the biradical region strongly depend on the CAS size. For the multiple bond changes in this pathway, inclusion of the O2p orbital in the active space or not determines what path, i.e. M or 20, actually exists for the given level of theory. At the (10,10) CAS level, putting the O2p orbital in the MCSCF core leads to path M with a symmetric, A′ TS. Starting from this TS geometry and placing the O2p in the active space for a (12,11) or (12,12) CAS leads to an asymmetric TS that leads to path 20 from Nguyen et al. [6] rather than path M.

The complexity with active space continues when examining the IRCs. For pathway M with the (10,10) CAS, ORF occurs towards the end of the IRC for the products. This ORF occurs during CO π bond formation, which should be obvious given the product geometries. Thus, since a CAS with the O2p is required for the products and does not give the TS for pathway M, the reverse barrier is unavailable at the MRMP2 level, and since there is no viable IRC, the CR-CC(2,3) barrier is also not reported. With the (10,10) CAS at the MRMP2 level, the forward barrier is 0.3 kcal/mol. Shifted CR-CC(2,3) values give the forward barrier as -2.7 kcal/mol.

Pathway 20 is found by either a (12,11) or (12,12) CAS and is the dissociation of a hydrogen molecule from acetaldehyde. It is the second of the pathways in the series of compromised active spaces due to computational expense, with pathway 13 being the first. To fully describe all of these compromised pathways requires at least a (16,15) or (16,16) CAS if no ORF occurs. For pathway 20 the compromised (12,12) CAS comes from taking acetaldehyde's methyl group and placing two of the three initially degenerate CH σs in the core. For the IRC leading to acetaldehyde using a compromised CAS, the active CH σ, σ* orbitals do not lie in the plane of symmetry. As well, when the IRC reaches the $H_2CCO + H_2$

products separated by ~4 Å, the bimolecular product does not fully separate, and H_2 does not lie in the symmetry plane of the H_2CCO molecule. Therefore, symmetry was not used in these calculations. Also, this particular pathway forms another CO π, π* and thus has IOC for the $H_2CCO + H_2$ ~4 Å product. The single point MRMP2 shifts the minima along the reaction coordinate very little, and the forward barrier (reverse barrier) is 76.8 (52.8) kcal/mol at the MRMP2 level. The experimental reaction enthalpy is 27.8 kcal/mol[58,60-63], which lies within 4 kcal/mol of the MRMP2 enthalpy. The CR-CC(2,3) (CR-CC(2,3) without MCSCF ZPE) level results in a forward barrier of 81.4 (54.4) kcal/mol, a reverse barrier of 87.4 (52.1) kcal/mol, and an overall enthalpy of 27.0 kcal/mol, which lies within 1 kcal/mol of the experimental enthalpy.

For path 20 in the forward barrier, the HH distance goes from 2.58 to 1.03 Å while for the reverse barrier, it changes from 1.03 to 0.76 Å. The methyl group CH distance changes from 1.11 to 1.68 Å for the forward barrier. The CC distances are 1.53, 1.42, and 1.33 Å at the stationary points in this pathway. This study's TS has only a slight difference from Nguyen's TS; the departing hydrogen from the terminal carbon lies 1.51 (1.38) Å from the carbon nearest the oxygen in this (Nguyen's) study. (The Supporting Information contains plots of both paths M and 20.)

Geometry Comparisons between the $\bullet CH_2CH_2O\bullet$ Biradical and TSs 10 – 13:

Despite the different active spaces, TSs 10 – 13 display several geometrical similarities. The smallest, OCCH dihedral angle ranges from 10 ^{o} to 24 ^{o}, and the OCC angle ranges from 115 ^{o} to 125 ^{o}. The CO distance varies from 1.34 to 1.41 Å, and the CC distance has a range of 1.43 to 1.51 Å. For the non-action Hs closest to the O, the two OCH angles cannot be distinguished. Likewise, for the Hs farthest from the O, the OCH angles also show

no substantial differences (e.g. range from 116 ^{o} to 120^{o}). Indeed, besides slight differences in the action areas that differentiate key features of the reactions themselves, these four TSs appear almost identical. Furthermore, these TS geometries from paths 10 – 12 differ little from the •CH_2CH_2O• biradical geometries obtained with the various CAS spaces: the CO distances differ by less than 1.00 Å, the CC distances also differ by less than 1.00 Å, and the OCCH dihedral (OCC angle) differs by 18 ^{o} (9 ^{o}) at most.

Surface Crossings at the MCSCF Level: •CH_2CH_2O• and •CH_2•CHOH

This study found two relevant surface crossings – one near the •CH_2CH_2O• biradical and one near the •CH_2•CHOH biradical. GAMESS currently cannot use MRMP2 for the MEX search because the MEX search requires analytic gradients; so MCSCF wavefunctions are used in these MEX calculations. Therefore, it is important to find an active space that gives consistent results with expected correlation effects. In order to justify the active space choice for a MEX of interest, Table 4 shows CAS and MRMP2 singlet-triplet energy gaps at •CH_2CH_2O• for the (6,6) and (8,8) CAS pure electronic states. The (6,6) CAS for the MEX and path 1 are the same. These singlet-triplet gaps show how increasing the active space by adding the lone O2p leads to inconsistent correlation recovery in the MRMP2 results relative to the CASSCF results; i.e., single state MRMP2 fails when the active space leads to quasi-degenerate electronic states (this perturbative, single state root flipping is well known[22]).

The •CH_2CH_2O• → •CH_2•CHOH reaction on the triplet surface (i.e. pathway 6 from our first study [1]) contains both crossings. The •CH_2CH_2O• MEX geometry has a CO distance of 1.43 Å and a small OCCH dihedral of 61 ^{o} whereas the •CH_2•CHOH MEX occurs at a CO distance of 1.40 Å and a small OCCH dihedral of 67 ^{o}. Paths 1 and 6 SOC

results and further investigation into the $\bullet CH_2CH_2O\bullet$ biradical give better insight into the PES near these MEXs.

SOC single point energies: $\bullet\bullet O + C_2H_4 \rightarrow \bullet CH_2CH_2O\bullet$

To further understand the multiple, close-lying PESs in the reaction, additional calculations were performed along the IRC for path 1. As previously discussed, since a pure state, (8,7) CAS has convergence difficulties for path 1, CASSCF optimizations were done with a (6,6) CAS. However, the (6,6) CAS does not allow for simultaneous exploration of all close-lying states. The (8,7) SA-CAS contains the lone O2p orbital, which essentially allows for the important configurations necessary for two of the four quasi-degenerate states near $\bullet CH_2CH_2O\bullet$. As discussed in the Methods section, (8,7) SA-CAS, single point energies followed by SO-MCQDPT2, single point energies should give the quasi-degenerate states and perturbatively recover the dynamic correlation and SOC effects from those several close-lying states.

Calculations using the (8,8) SA-CAS, MCQDPT2 method were also performed for the biradical region to see if it would be an effective CAS size. The (8,8) SA-CAS configurations qualitatively differ from the (8,8) MCQDPT2 configurations whereas configurations between the (8,7) SA-CAS and (8,7) MCQDPT2 do not qualitatively differ. In other words, for the (8,8) SA-CAS, the dominant configurations drastically change with the perturbation. Thus, this study uses an (8,7) SA-CAS reference to further examine path 1 as well as the $\bullet CH_2CH_2O\bullet$ biradical region rather than an (8,8) CAS.

For pathway 1 the nonrelativistic, triplet A″ IRC from the O + ethylene to the $\bullet CH_2CH_2O\bullet$ biradical has the nonrelativistic, triplet A′ state lying very close throughout the path. To obtain the relativistic potentials, (8,7) SA-CAS, single point energies followed by

SO-MCQDPT2, single point energies were calculated at the (6,6) CAS IRC geometries. However, for the SA-CAS, only careful state selection yields physically reasonable, smooth curves. Using an (8,7) SA-CAS with the three lowest-lying, nonrelativistic triplets gives a reasonable wavefunction for obtaining SOC results – nine relativistic triplets and five relativistic singlets – until near the triplet A″ TS (see Figure 3 - 5). Since SO-MCQDPT2 and SO-CASCI coupling constants are in good agreement, only the latter are given.

The three nonrelativistic triplets no longer remain degenerate near the $\bullet CH_2CH_2O\bullet$ biradical. One of these surfaces becomes much higher in energy and prevents correct results with MCQDPT2 in that surface region (see Supporting Information for these additional graphs). Because of this limitation, additional calculations were repeated with an (8,7) SA-CAS using the two lowest-lying, nonrelativistic triplets and the two lowest-lying, nonrelativistic singlets in order to obtain SOC results across the entire IRC for path 1. Figs. 6 - 8 display the results from the TS to $\bullet CH_2CH_2O\bullet$. In particular, Figure 8 shows a sharp but continuous change in the coupling constants; this change corresponds to the OCCH dihedral rotation at the end of path 1. Thus, these four states (two triplets and two singlets) will prove very important in determining any dynamics for this system.

SOC single point energies: $\bullet CH_2CH_2O\bullet \rightarrow \bullet CH_2\bullet CHOH$

Pathway 6 on the lowest-lying triplet surface has quasi-degeneracy with several surfaces. In this case, four states lie near the $\bullet CH_2CH_2O\bullet$ biradical, but only a triplet and singlet lie close between the TS and $\bullet CH_2\bullet CHOH$. Using the (8,8) CAS IRC geometries, several SA-CAS wavefunctions were examined to give more information on the close-lying states for this part of the PES. For example, a (10,9) SA-CAS would be appropriate to describe all fours states. However, ORF does not allow the SA-CAS to be consistent at all

geometries: the O2p* IOC orbital displaces the relevant CH σ* orbital despite various attempts to maintain the active space. Therefore, an (8,8) SA-CAS is used to obtain accurate SOC results for the lowest-lying triplet and close-lying singlet for the full path as compared to the lowest state results from the (10,9) SA-CAS on the $\bullet CH_2CH_2O\bullet$ biradical half of the path. Figure 9 shows the energetics for this hydrogen migration at the (8,8) SO-CASCI level. All coupling constants along this path are less than 4 cm^{-1}. These SOC constants do not match the (10,9) CAS couplings near $\bullet CH_2CH_2O\bullet$ since the (8,8) CAS does not contain some states of interest due to its lack of the O2p in the active space.

SOC single point energies around the $\bullet CH_2CH_2O\bullet$ biradical:

With several singlet TSs in this PES region (TS 10 – 13) and several close-lying surfaces, a finer description of the $\bullet CH_2CH_2O\bullet$ biradical region of the PES requires two nonrelativistic singlet and two nonrelativistic triplet states. Path 1 leads to large changes in the OCCH dihedral near the $\bullet CH_2CH_2O\bullet$ biradical. Thus, in order to describe these lower state curvatures over an extended geometry range, (6,6) CAS triplet constrained optimizations and (6,6) CAS singlet single point energies were performed on a grid from 1.2 – 4 Å for the CO distance and from 0 – 90 ^{o} for the OCCH dihedral in increments of 0.5 Å and 15 ^{o}, respectively. The boundary of 1.2 Å was chosen since ORF occurs towards the more repulsive part of the potential. Figure 10 shows this region of the PES from these calculations. A crude outline of the MEP of path 1 can be seen along the 90 ^{o} dihedral. As the initial oxygen approach reaches an OCCH dihedral angle of 0 ^{o}, the constrained optimizations no longer lead to the biradical but to the initial minimum for hydrogen abstraction from ethylene.

In order to refine the results in the quasi-degenerate PES region, the same calculations were performed as above, but with a narrower range of geometries and a finer resolution; i.e. for 1.2 – 2 Å and 50 – 90 ^{o} in increments of 0.05 Å and 1 ^{o}, respectively. Then, (8,7) SA-CAS (see Figure 11) and SOC calculations were performed on the four lowest-lying states as described above. As shown in Figure 11, these lowest-lying states are nearly degenerate at many points around the biradical region. Figs. 12 - 13 give the CASCI and MCQDPT2 contours of the lowest-lying, nonrelativistic singlet relative to the lowest-lying, nonrelativistic triplet. Figure 14 displays the SOC coupling constants for the lowest-lying, singlet-triplet gap in this region (the Supporting Information contains excited state plots of this region). In particular, this figure shows that the relativistic energy changes as the OCCH dihedral rotates. The coupling constants vary in a 'half cone-like' fashion with respect to the OCCH dihedral angle near ~1.4 Å CO distance and ~60 ^{o} OCCH dihedral – coordinates near the (6,6) CAS, $\bullet CH_2CH_2O\bullet$ MEX. It is expected that this PES region will strongly influence the non-adiabatic dynamics as a result of the many quasi-degenerate states, showing that a single state, single reference approach may miss much of the essential non-adiabaticity.

Conclusions:

In this work with careful choice of the active space, CASSCF, MRMP2, MR-CISD, and MR-AQCC are used to model the multiconfigurational nature of the singlet surface for the addition of oxygen atom to ethylene. Reaction paths involving the biradical region (paths 1 and 10 – 13) clearly have many configurations contributing to the wavefunction and thus require a multiconfigurational treatment to account for non-dynamical correlation. In particular, with pure state solutions, TSs 10-13 all have similar geometries with MEPs in a very flat PES region. Relative to available data from single reference treatments,

multireference calculations resulted in the following significant differences for TS geometries in this region: 1) for path 13 the CO distance with the (10,10) CAS results in 1.40 Å while Nguyen's[6] B3LYP yields 1.31 Å, and 2) for path 20 for one action CH distance, (12,12) CAS shows 1.51 Å whereas Nguyen's[6] B3LYP gives 1.38 Å.

While perturbative, dynamic correlation does not have much effect on the geometries of the singlet stationary points, dynamic correlation is *essential* for obtaining reasonable energetics. Although the MRMP2 captures a significant portion of the dynamic correlation, this treatment certainly does not encapsulate all of it. While multiple levels of theory have been used in this study, the best level of theory for barrier energetics is still not well understood. Since the biradical regions clearly require multiconfigurational wavefunctions, it is anticipated that CR-CC(2,3) barriers may require more nondynamical correlation while the MRMP2 barriers require more dynamic correlation. The MR-AQCC calculations hold promise, but are limited to smaller basis sets and active spaces because of the computational expense. CR-CC(2,3) gives values within 1 kcal/mol of the experimental values for the thermodynamics of pathways 13 and 20, while MRMP2 is within 1.5 kcal/mol for pathway 13 and within about 4.0 kcal/mol for pathway 20. Computational expense limits the latter pathway to a compromised CAS, which may explain a significant part of this error. However, these comparisons with experiment are for closed shell molecules and do not help to elucidate the best methods for obtaining energetics in the biradical regions.

At the MRMP2 and CR-CC(2,3) levels, pathways 1, 10, 11, 12, and M are essentially barrierless processes, which again emphasizes the need for a careful treatment of this part of the PES. Furthermore, in paths 13 and M, the addition of O2p to the active space qualitatively changes these pathways at the CASSCF level in the sense that a geometrically

similar TS leads to different minima. However, whenever possible, CASSCF active spaces for pure states do not contain lone O2p orbitals, which tend to lead to IOC ORF with the current, convergence techniques. This active space construction is particularly important in both solving for pure states at the CASSCF level (avoiding IOC ORF) and obtaining single state MRMP2 results (avoiding state root flipping).

In addition to the pure state calculations of the reactions, this study explicitly and appropriately examines the biradical regions. First, two, separate CASSCF surface crossings between the lowest-lying triplets and lowest-lying singlets are determined. Second, an (8,7) SA-CAS is used to describe the $\bullet CH_2CH_2O\bullet$ biradical around this CASSCF approximation to the given surface crossing. Third, SO-MCQDPT2 is performed on the lowest four states in a region around $\bullet CH_2CH_2O\bullet$. Without treatments such as non-orthogonal MCSCF, these quasi-degenerate states currently require SA wavefunctions to sort out close-lying states – with or without IOC. This active space selection requires care since MCQDPT2 application to an (8,8) SA-CAS reference changes the dominant configuration(s). This fact also supports the avoidance of improper IOC in SA-MCSCF, reference wavefunctions.

Overall, the O + ethylene PES definitely requires a complex treatment in the quasi-degenerate regions around the MEXs. In fact, PES regions about paths 10 - 13 will need more sophisticated, multireference methods incorporating dynamic correlation in order to develop an accurate potential for dynamic simulations. Future work on O + ethylene will take a number of directions that include further GVB-PP and MR-AQCC energetics, various coupled cluster calculations, and dynamical simulations including but not limited to surface hopping.

Acknowledgments:

The authors are indebted to Michael W. Schmidt and Mark S. Gordon for help in using the capabilities of GAMESS and MCSCF. This material is based upon work supported by the National Science Foundation under Grant No. OISE-0730114 for the Partnerships in International Research and Education (PIRE) and by the Robert A. Welch Foundation under Grant No. D-0005. TeraGrid resources were provided by the Texas Advanced Computing Center (TACC). Support was also provided by the High-Performance Computing Center (HPCC) at Texas Tech University, under the direction of Philip W. Smith. In addition, this work was supported by the Austrian Science Fund within the framework of the Special Research Program F16 (Advanced Light Sources) and Project P18411-N19. Computer time at the Vienna Scientific Cluster (project no. 70019) is gratefully acknowledged. TLW acknowledges computing resources purchased through funds provided by Ames Laboratory and Iowa State University.

Electronic Supplemental Material:

The online version of this article (doi:10.1007/s00214-012-1123-0) contains supplementary material, which is available to authorized users.

Table 1. Description of the main CAS sizes used for each reaction pathway. Only the lone pairs and bonding orbitals are shown; the antibonding orbitals can be inferred from bonding orbitals and the indication of in-out correlation. The O2s is always in the core.

reaction pathway	active space	in-out correlation	description
$1\,^3A''$ $1\,^3A'$ $1\,^1A''$	(6,6)	N	CCσ, CCπ, biradical(2O2p) → CCσ, COσ, biradical(farC2p, O2p)
$1\,^1A_1$	(6,6)	N	CCσ, CCπ, biradical(2O2p) → CCσ, 2COσ
10	(14,14)	N	4 CHσ, CCσ, COσ, biradical(farC2p, O2p) → 4 CHσ, CCσ, COσ, COπ
11	(6,6)	N	COσ, CCσ, biradical(farC2p, O2p) → 2 COσ, CCσ
12	(10,10)	N	COσ, CCσ, 2 CHσ, biradical(farC2p, O2p) → COσ, CCσ, CHσ, OHσ, CCπ
13	(10,10)	Y	CHσ, COσ, COπ, CCσ, lone O2p → CHσ, 2 COσ, CCσ, lone O2p
M	(10,10)	N/A	2 CHσ, COσ, CCσ, biradical(farC2p, O2p) → ORF for the O2p
20	(12,12)	Y	2 CHσ, COπ, lone O2p, COσ, CCσ → HHσ, 2 COπ, CCπ, COσ, CCσ,

Table 2. Barriers to reactions in kcal/mol calculated at the CAS level without and with (in parenthesis) ZPE. Nguyen's barriers include ZPE.

reaction pathway	CAS forward barrier[a]	CAS reverse barrier[a]	Nguyen forward barrier[b]	Nguyen reverse barrier[b]
Triplet A′ 1	19.8 (19.6)	17.4[c]	-	20.9
Singlet A_1 1	2.2 (2.5)	108.9 (104.8)	-	-
Singlet A″ 1	5.2 (5.2)	No Cs Minimum	-	-
10	1.5 (1.0)	90.5 (86.6)	-2.9	82.0
11	0.2 (0.3)	59.8 (55.7)	1.6	59.1
12	6.8 (6.3)	80.0 (75.5)	0.0	74.2
13	95.0 (92.0)	64.8 (60.5)	86.3	58.9
M	17.1 (14.9)	ORF	-	-
20	95.0 (89.0)	66.6 (68.9)	81.3	55.6

a. CAS barriers are obtained from unconstrained optimizations unimolecular species and a distance of ~3-5 Å for separated products.

b. Nguyen's barriers are the average barriers from several different levels of theory.[6]

c. This difference does not contain ZPE and is not a true barrier since it is taken from a Cs geometry that is not a minimum.

Table 3. Barriers to reactions in kcal/mol calculated at the MRMP2 and CR-CC(2,3) levels with aug-cc-pVTZ basis. The MR-AQCC barriers are calculated with the cc-pVDZ basis. F indicates the forward barrier while R indicates the reverse barrier. Absence of F or R indicates reaction energies rather than barriers. The shifted MRMP2 barriers are formed based on the minima and maxima values of the MRMP2 single point energies along the CASSCF IRC. Stationary point (SP) MRMP2 and shifted MRMP2 barriers include ZPE from CASSCF stationary points while MR-AQCC barriers derive from stationary points and ZPE at the MR-AQCC level. Nguyen's barriers include ZPE. Vertical MRMP2, vertical MR-AQCC, and vertical CR-CC(2,3) barriers do not include ZPE and are all at the same CASSCF geometries.

reaction pathway	SPMRMP2[a]	Shifted MRMP2	MR-AQCC	Nguyen [b]/Yang	Vertical MRMP2	Vertical MR-AQCC	Vertical CR-CC(2,3)
Triplet A′ 1 R	17.6[j]	21.2[j]	-	20.9	-	-	-
Triplet A′ 1 F	-0.8	2.8 3.3[c]	-	-	-0.6	-	-
Singlet A_1 1	-121.5	-	-	-	-	-	-
Singlet A″ 1	-61.7[d]	-	-	-	-	-	-
10	-77.6	-	-	-84.9	-81.0	-	-93.3
11 R	56.6	56.6	-	59.1	60.6	-	68.6
11 F	0.0	0.0 0.2[c]	-	1.6	0.0	-	-0.3
12 R	64.1	64.1	69.7[g,h]	74.2	67.0	74.6[h]	75.9 73.5[i]

12 F	-3.4	-3.4 0.4[c]	1.8[g,h]	0.0	-2.9	3.4[h]	-6.3 -4.4[i]
13 R	52.5	52.7	51.7[h]	58.9 53.9[e]	56.9	55.8[h]	68.1 63.0[i]
13 F	78.3	78.5	83.0h	86.3 81.8[e]	81.3	85.9[h]	94.2 90.8[i]
M R	ORF[f]	ORF[f]	-	-	ORF[f]	-	ORF[f]
M F	-1.6	-1.1 0.3[c]	-	-	0.6	-	-6.7
20 R	52.6	52.6 52.8[c]	-	55.6	50.4	-	52.1
20 F	76.7	76.8	-	81.3	82.8	-	87.4

a. SPMRMP2 values derive from CAS stationary points, which are obtained from unconstrained optimizations of unimolecular species and a distance of ~3-5 Å for separated products.

b. Nguyen's barriers are the average barriers from several different levels of theory.[6]

c. MRMP2 barrier with ZPE from shifts in minima in addition to shifts in TSs.

d. Barrier does not include ZPE.

e. Yang's barriers at the CCSD(T)/cc-pVTZ//B3LYP/6-311++G* level of theory.[14]

f. ORF indicates orbital root flipping (see text for more complete description).

g. This TS contains an extra imaginary frequency, which could not be eliminated for pathway 12 to retain an actual barrier.

h. Barrier is with cc-pVDZ basis.

i. Barrier is with aug-cc-pVDZ basis.

j. This difference does not contain ZPE and is not a true barrier since it is taken from a Cs geometry that is not a minimum.

Table 4. (6,6) and (8,8) CAS and MRMP2 triplet-singlet gaps in kcal/mol at the $\bullet CH_2CH_2O\bullet$ biradical triplet minimum.

Triplet-Singlet Gap (kcal/mol)	(6,6) Triplet Minimum Geometry	(8,8) Triplet Minimum Geometry
(6,6) CAS	0.6	0.8
(6,6) MRMP2	0.7	0.9
(8,8) CAS	0.4	1.3
(8,8) MRMP2	-0.4	0.5

Figure 1. TS geometry at the (14,14) CAS level of theory shown for Pathway 10.

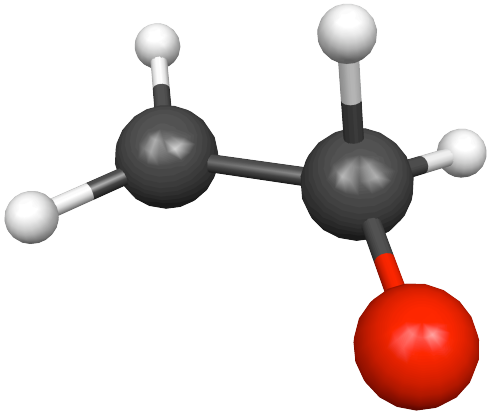


Figure 2. (14,14) CAS IRC and MRMP2 single point energies for Pathway 10 shown with shifted energies so that the CAS energy equals the corresponding MRMP2 and CR-CC(2,3), single point energies at the •CH_2CH_2O• biradical. These results show how dynamic correlation leads to a barrierless reaction.

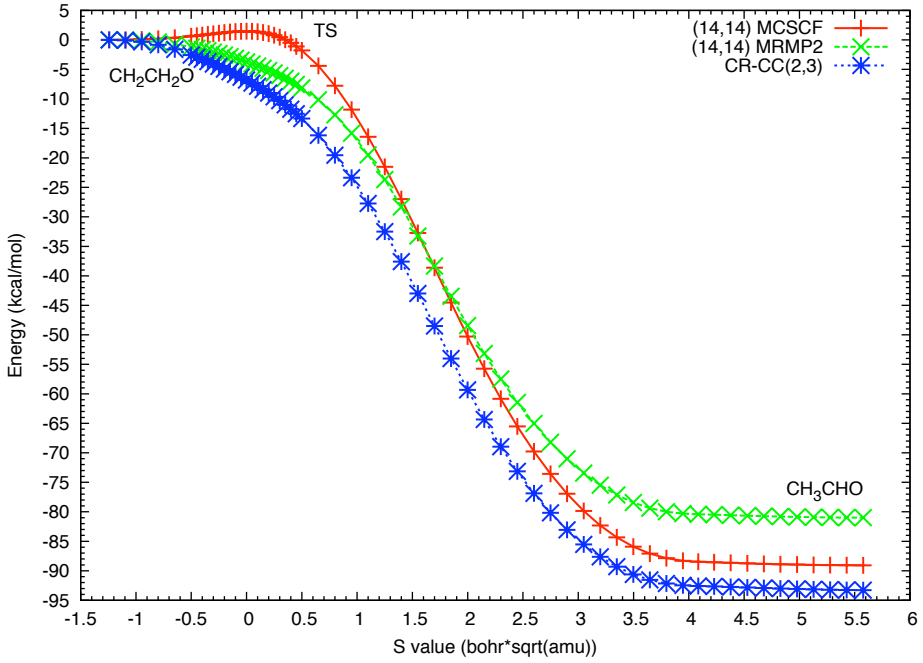


Figure 3. Path 1 SO-CASCI single point energies shown with nine triplets and five singlets from an (8,7) SA-CAS on the three lowest-lying, nonrelativistic triplets. Here, the following pairs of relativistic state energies almost always overlap: singlets 4 and 5; and the upper 3 triplets.

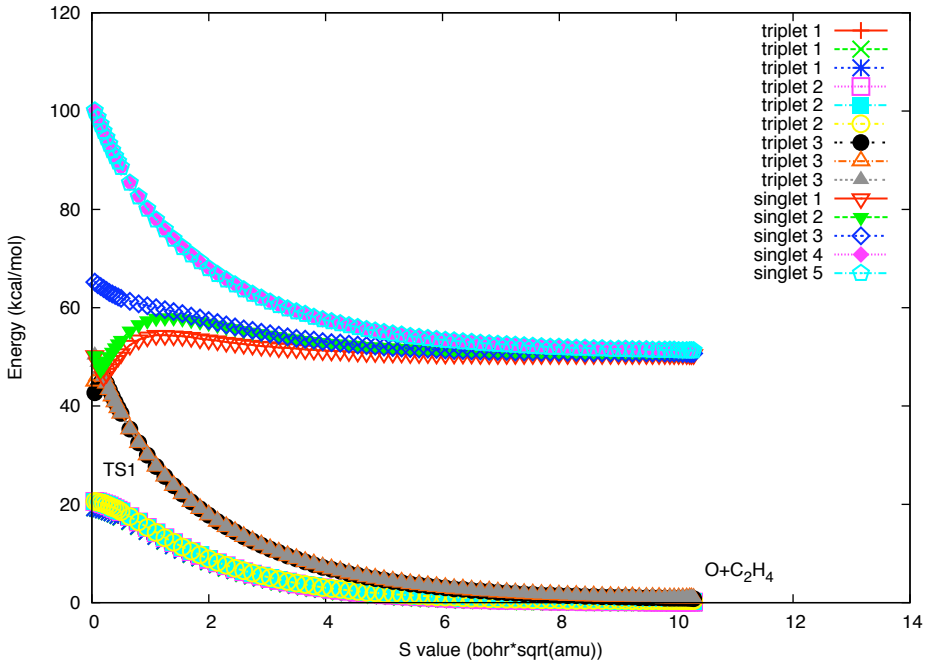


Figure 4. Path 1 SO-MCQDPT2 single point energies shown with nine triplets and five singlets from an (8,7) SA-CAS on the three lowest-lying, nonrelativistic triplets. . Here, the following pairs of relativistic state energies almost always overlap: singlets 4 and 5; and the upper 3 triplets.

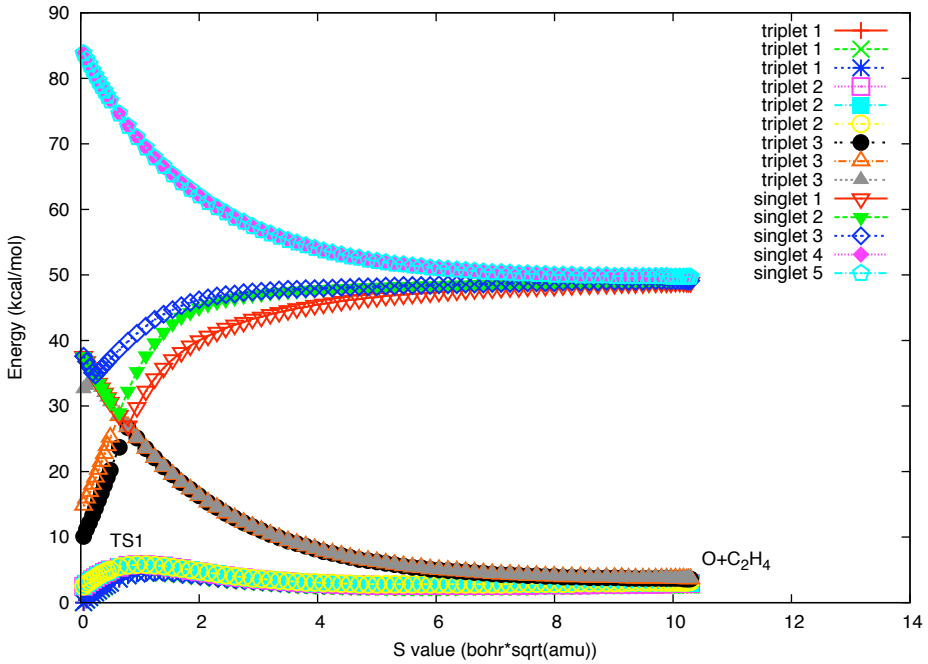


Figure 5. Path 1 SO-CASCI single point coupling constants shown with nine triplets and five singlets from an (8,7) SA-CAS on the three lowest-lying, nonrelativistic triplets. Here, the following sets of pairs of relativistic state energies almost always overlap: triplets 1 and 3 with triplets 2 and 3; singlet 4 and triplet 1 with singlet 5 and triplet 2; and pairs for singlet 2 and triplet 1, singlet 4 and triplet 3, and singlet 5 and triplet 3.

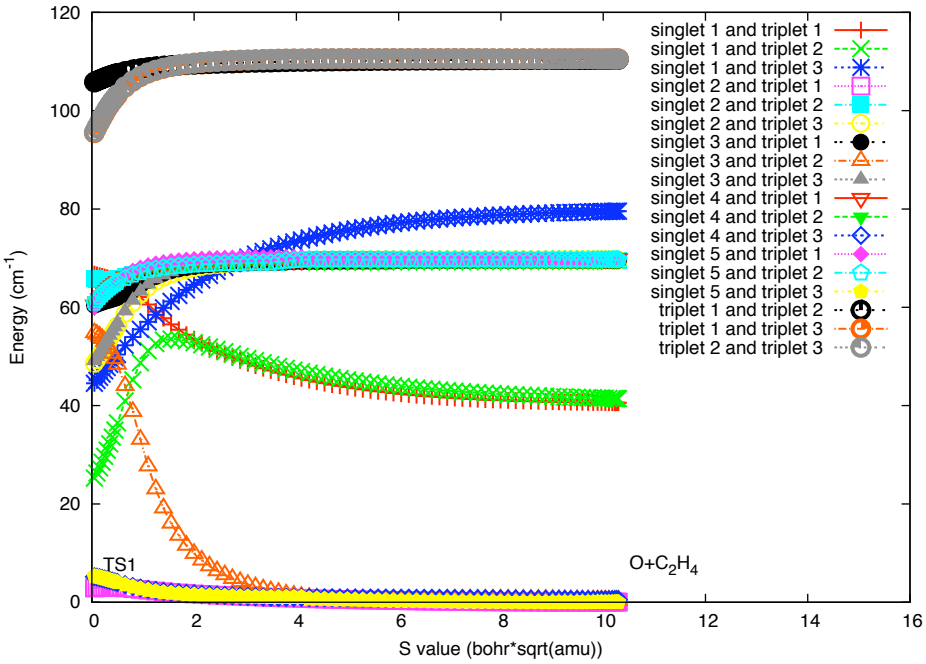


Figure 6. Path 1 SO-CASCI single point energies shown with six triplets and two singlets from an (8,7) SA-CAS on the two lowest-lying, nonrelativistic triplets and two lowest-lying, nonrelativistic singlets. Here, the following relativistic state energies almost always overlap: the first 3 triplets; and the second 3 triplets.

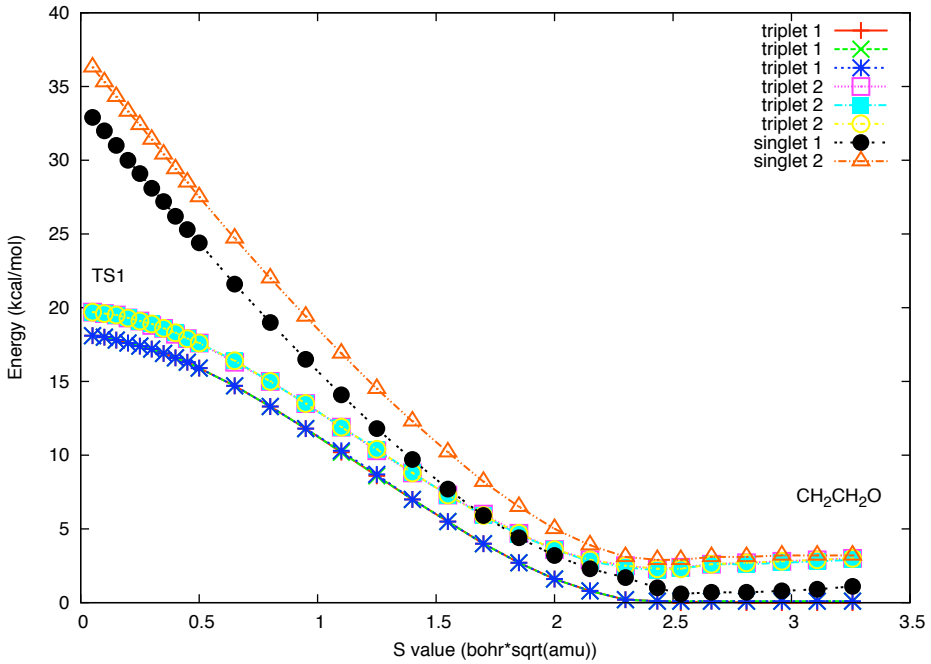


Figure 7. Path 1 SO-MCQDPT2 single point energies shown with six triplets and two singlets from an (8,7) SA-CAS on the two lowest-lying, nonrelativistic triplets and two lowest-lying, nonrelativistic singlets. Here, the following relativistic state energies almost always overlap: the first 3 triplets; and the second 3 triplets.

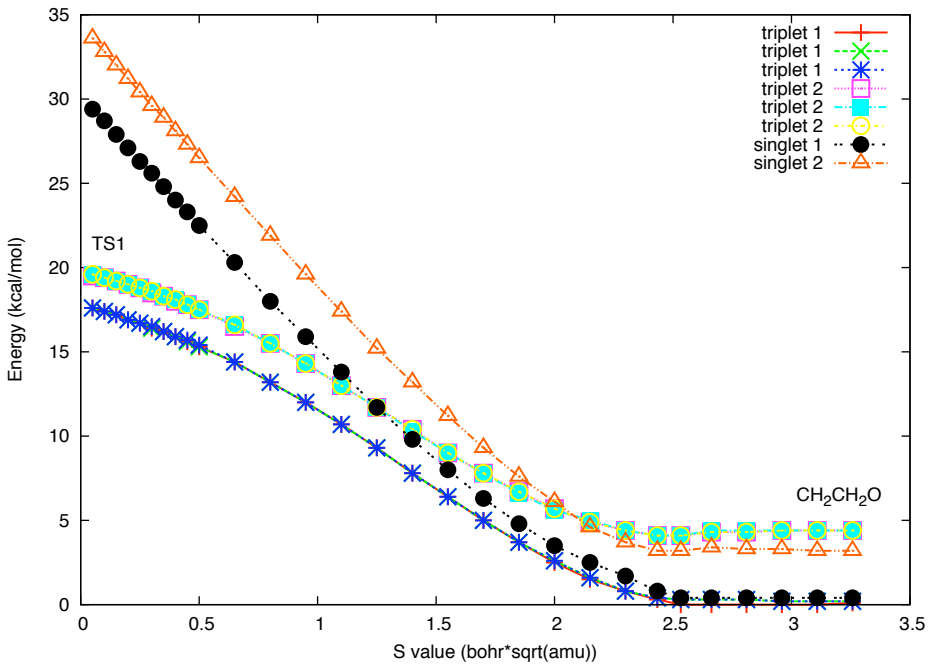


Figure 8. Path 1 SO-CASCI single point, coupling constants shown with six triplets and two singlets from an (8,7) SA-CAS on the two lowest-lying, nonrelativistic triplets and two lowest-lying, nonrelativistic singlets.

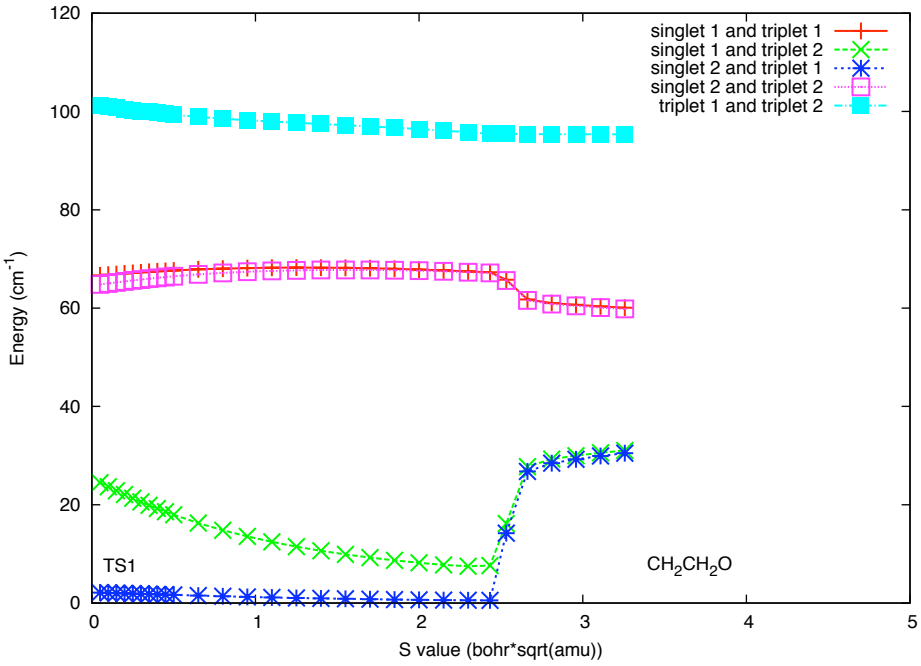


Figure 9. Path 6 SO-CASCI single point energies shown with three relativistic triplets and one singlet from an (8,8) CAS that is SA with the lowest-lying, nonrelativistic triplet and lowest-lying, nonrelativistic singlet. Here, the three relativistic triplet curves are indistinguishable because of overlap in the energies.

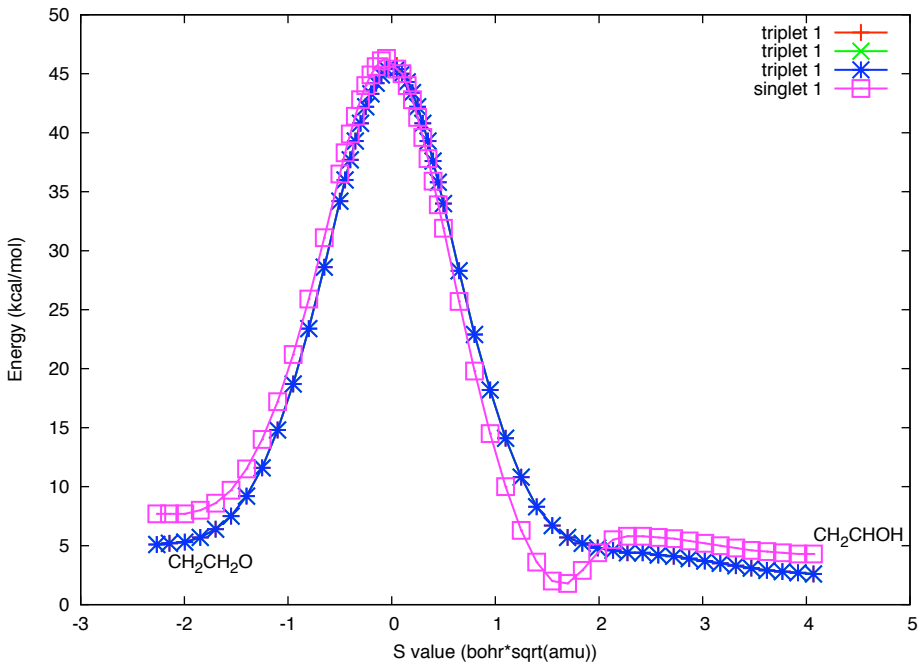


Figure 10. (6,6) CAS triplet constrained optimizations and (6,6) CAS singlet single point energies on a large region of PES for overview energetics.

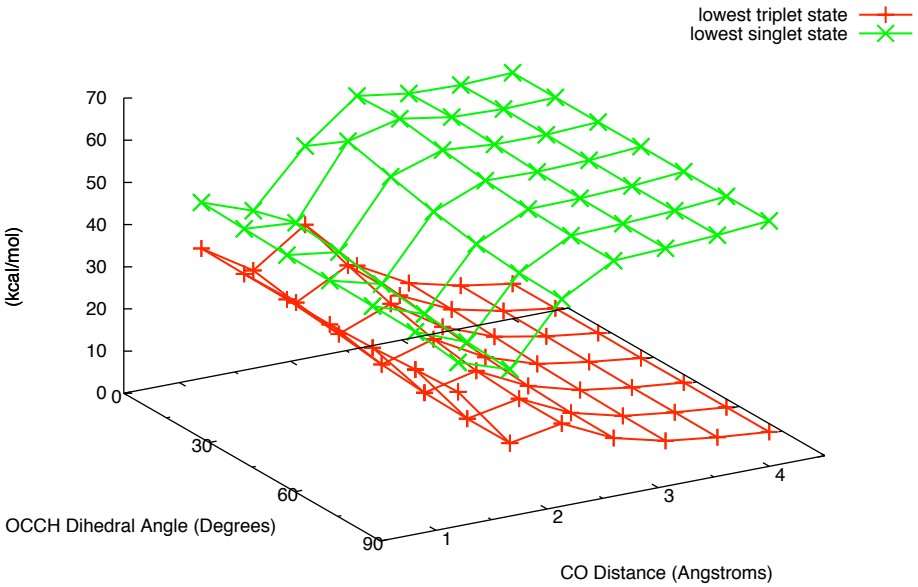


Figure 11. (8,7) SA-CAS single point energies for 4 lowest-lying states – two singlets and two triplets – on top of (6,6) CAS triplet constrained optimizations about the •CH_2CH_2O• biradical.

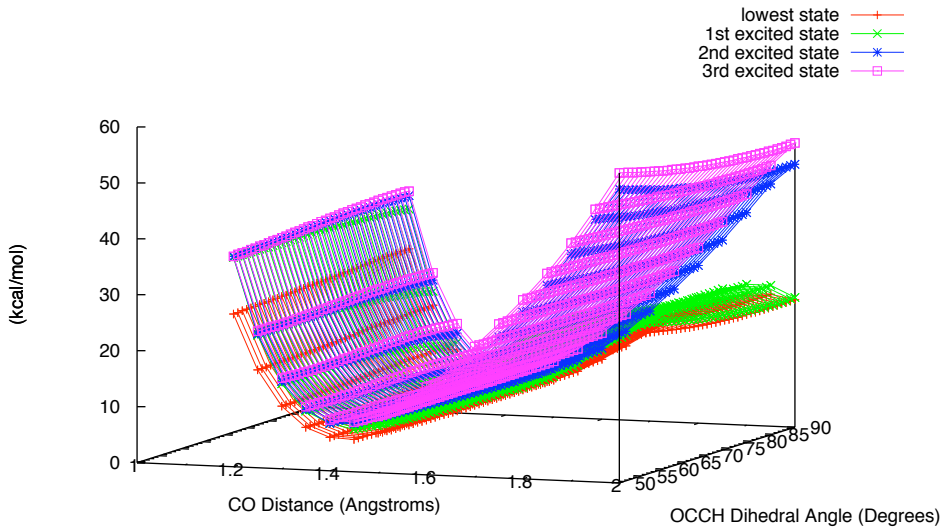


Figure 12. CASCI single point energies from an (8,7) SA-CAS in kcal/mol for lowest-lying, singlet-triplet gap about the •CH_2CH_2O• biradical.

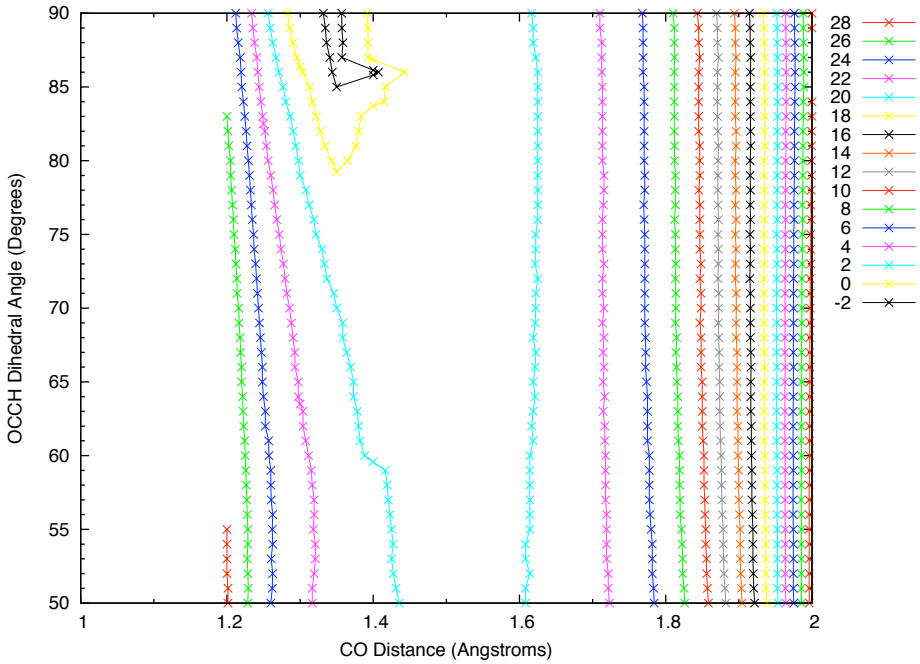


Figure 13. MCQDPT2 single point energies from an (8,7) SA-CAS in kcal/mol for lowest-lying, singlet-triplet gap about the •CH_2CH_2O• biradical.

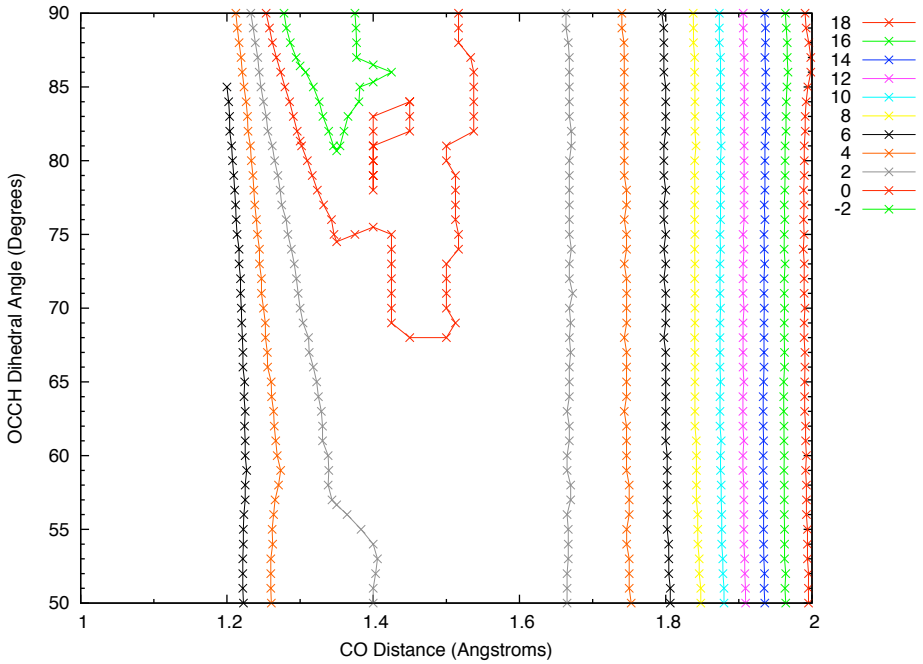
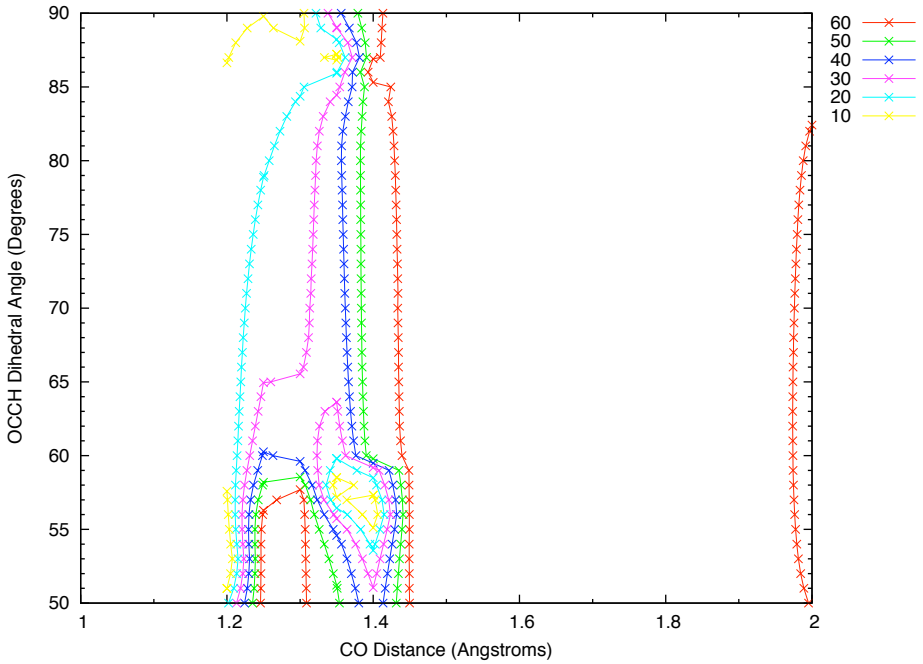


Figure 14. SO-CASCI single point coupling constants from an (8,7) SA-CAS in cm^{-1} for lowest-lying, singlet-triplet gap about the $\bullet CH_2CH_2O\bullet$ biradical.



References:

[1]West, A. C.; Kretchmer, J. S.; Sellner, B.; Park, K.; Hase, W. L.; Lischka, H.; Windus, T. L. *J. Phys. Chem. A* **2009**, *113*, 12663.

[2]Gardiner, W. C., Jr.; Editor *Combustion Chemistry*; Springer-Verlag, New York, 1984.

[3]Casavecchia, P.; Capozza, G.; Segoloni, E.; Leonori, F.; Balucani, N.; Volpi Gian, G. *J Phys Chem A* **2005**, *109*, 3527.

[4]Su, H.; Zhao, S.; Liu, K.; Xiang, T. *J Phys Chem A* **2007**, *111*, 9600.

[5]Lee, S.-H.; Huang, W.-J.; Chen, W.-K. *Chem. Phys. Lett.* **2007**, *446*, 276.

[6]Nguyen, T. L.; Vereecken, L.; Hou, X. J.; Nguyen, M. T.; Peeters, J. *J. Phys. Chem. A* **2005**, *109*, 7489.

[7]Yamaguchi, K.; Yabushita, S.; Fueno, T.; Kato, S.; Morokuma, K. *Chem. Phys. Lett.* **1980**, *70*, 27.

[8]Dupuis, M.; Wendoloski, J. J.; Takada, T.; Lester, W. A., Jr. *J. Chem. Phys.* **1982**, *76*, 481.

[9]Fueno, T.; Takahara, Y.; Yamaguchi, K. *Chem. Phys. Lett.* **1990**, *167*, 291.

[10]Smith, B. J.; Nguyen Minh, T.; Bouma, W. J.; Radom, L. *J. Am. Chem. Soc.* **1991**, *113*, 6452.

[11]Jursic, B. S. *THEOCHEM* **1999**, *492*, 85.

[12]Hu, W.; Lendvay, G.; Maiti, B.; Schatz, G. C. *J. Phys. Chem. A* **2008**, *112*, 2093.

[13]Joshi, A.; You, X.; Barckholtz Timothy, A.; Wang, H. *J Phys Chem A* **2005**, *109*, 8016.

[14]Yang, X.; Maeda, S.; Ohno, K. *J. Phys. Chem. A* **2007**, *111*, 5099.

[15]Shepler, B. C.; Braams, B. J.; Bowman, J. M. *J. Phys. Chem. A* **2008**, *112*, 9344.

[16]Heazlewood, B. R.; Jordan, M. J. T.; Kable, S. H.; Selby, T. M.; Osborn, D. L.; Shepler, B. C.; Braams, B.; Bowman, J. M. *Proc. Natl. Acad. Sci. U. S. A., Early Ed.* **2008**, 1.

[17]Shepler, B. C.; Braams, B. J.; Bowman, J. M. *J. Phys. Chem. A* **2007**, *111*, 8282.

[18]Miyoshi, A.; Yoshida, J.-i.; Shiki, N.; Koshi, M.; Matsui, H. *Phys. Chem. Chem. Phys.* **2009**, *11*, 7318.

[19]Gordon, M. S.; Schmidt, M. W. *Theory Appl. Comput. Chem.: First Forty Years* **2005**, 1167.

[20]Roos, B. *Adv. Chem. Phys.* **1987**, *69*, 399.

[21]Ruedenberg, K.; Sundberg, K. R. *Quantum Sci.* **1976**, 505.

[22]Schmidt, M. W.; Gordon, M. S. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233.

[23]Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

[24]Ivanic, J.; Ruedenberg, K. *Theor. Chem. Acc.* **2001**, *106*, 339.

[25]Lengsfield, B. H., III. *J. Chem. Phys.* **1980**, *73*, 382.

[26]Fletcher, G. D. *Mol. Phys.* **2007**, *105*, 2971.

[27]Yarkony, D. R. *Chem. Phys. Lett.* **1981**, *77*, 634.

[28]Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1980**, *72*, 880.

[29]Boys, S. F. *Quantum Theory At., Mol., Solid State. 1966* **1966**, 253.

[30]Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154.

[31]Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.

[32]Piecuch, P.; Kucharski, S. A.; Kowalski, K.; Musial, M. *Comput. Phys. Commun.* **2002**, *149*, 71.

[33]Piecuch, P.; Wloch, M. *J. Chem. Phys.* **2005**, *123*, 224105/1.

[34]Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481.

[35]Lischka, H.; Shepard, R.; Brown, F. B.; Shavitt, I. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1981**, *15*, 91.

[36]Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Muller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664.

[37]Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Mueller, T.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Boehm, H. J.; Chang, A.; Comeau, D. C.; Gdanitz, R.; Dachsel, H.; Ehrhardt, C.; Ernzerhof, M.; Hoechtl, P.; Irle, S.; Kedziora, G.; Kovar, T.; Parasuk, V.; Pepper, M. J. M.; Scharf, P.; Schiffer, H.; Schindler, M.; Schueler, M.; Seth, M.; Stahlberg, E. A.; Zhao, J.-G.; Yabushita, S.; Zhang, Z.; Barbatti, M.; Matsika, S.; Schuurmann, M.; Yarkony, D. R.; Brozell, S. R.; Beck, E. V.; Blaudeau, J.-P.; Ruckenbauer, M.; Sellner, B.; Plasser, F.; Szymczak, J. J. **2008**, www.univie.ac.at/columbus.

[38]Hay, P. J.; Hunt, W. J.; Goddard, W. A., III. *J. Amer. Chem. Soc.* **1972**, *94*, 8293.

[39]Hunt, W. J.; Hay, P. J.; Goddard, W. A., III. *J. Chem. Phys.* **1972**, *57*, 738.

[40]Shepard, R.; Kedziora, G. S.; Lischka, H.; Shavitt, I.; Mueller, T.; Szalay, P. G.; Kallay, M.; Seth, M. *Chem. Phys.* **2008**, *349*, 37.

[41]Bunge, A. V. *J. Chem. Phys.* **1970**, *53*, 20.

[42]Farazdel, A.; Dupuis, M. *J. Comput. Chem.* **1991**, *12*, 276.

[43]Manaa, M. R.; Yarkony, D. R. *J. Chem. Phys.* **1993**, *99*, 5251.

[44]Yarkony, D. R. *J. Phys. Chem.* **1993**, *97*, 4407.

[45]Witek, H. A.; Choe, Y.-K.; Finley, J. P.; Hirao, K. *J. Comput. Chem.* **2002**, *23*, 957.

[46]Kozlowski, P. M.; Davidson, E. R. *J. Chem. Phys.* **1994**, *100*, 3672.

[47]Malrieu, J.-P.; Heully, J.-L.; Zaitsevskii, A. *Theor. Chim. Acta* **1995**, *90*, 167.

[48]Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.

[49]Marian, C. M. *Rev. Comput. Chem.* **2001**, *17*, 99.

[50]Fedorov, D. G.; Koseki, S.; Schmidt, M. W.; Gordon, M. S. *Int. Rev. Phys. Chem.* **2003**, *22*, 551.

[51]Yarkony, D. R. *Int. Rev. Phys. Chem.* **1992**, *11*, 195.

[52]Furlani, T. R.; King, H. F. *J. Chem. Phys.* **1985**, *82*, 5577.

[53]King, H. F.; Furlani, T. R. *J. Comput. Chem.* **1988**, *9*, 771.

[54]Fedorov, D. G.; Gordon, M. S. *J. Chem. Phys.* **2000**, *112*, 5611.

[55]Langhoff, S. R.; Davidson, E. R. *Int. J. Quantum Chem.* **1974**, *8*, 61.

[56]Bruna, P. J.; Peyerimhoff, S. D. *Chem. Phys. Lett.* **1981**, *72*, 278.

[57]Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364.

[58]Chao, J.; Hall, K. R.; Marsh, K. N.; Wilhoit, R. C. *J. Phys. Chem. Ref. Data* **1986**, *15*, 1369.

[59]Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables, 4th ed.; J. Phys. Chem. Ref. Data, Mongr. 9; American Chemical Society: Washington, DC, 1998; p 1.*, *27*.

[60]Wiberg, K. B.; Crocker, L. S.; Morgan, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 3447.

[61]JANAF Thermochemical Tables; Natl. Stand. Ref. Data Ser.; U.S. National Bureau of Standards: Washington, D. **1985**, *37*.

[62]Ruscic B, B. J., Burcat A,Csaszar AG,Demaison J, Janoschek R, Martin JML, Morton ML, Rossi MJ, Stanton JF, Szalay PG, Westmoreland PR, Zabel F, Berces T "IUPAC Critical Evaluation of Thermochemical Properties of Selected Radicals. Part I" J. Phys. Chem. Ref. Data, Vol. 34, No. 2, 2005, 573.

[63]Cox, J. D.; Wagman, D. D.; Medvedev, V. A. *CODATA Key Values for Thermodynamics*, 1989.

CHAPTER 4. O + C_2H_4 POTENTIAL ENERGY SURFACE: LOWEST-LYING SINGLET AT THE MULTIREFERENCE LEVEL

Aaron C. West, Joseph D. Lynch, Bernhard Sellner, Hans Lischka, William L. Hase, and Theresa L. Windus

[Submitted to *Theor. Chem. Acc.*]

Abstract:

In previous studies[West, et.al., J Phys Chem A 113 (45):12663; West, et.al., Theor Chem Acc 131:1123] the lowest-lying O(^{3}P) + C_2H_4 and singlet PES near the •CH_2CH_2O• biradical were extensively explored at several levels of theory. In this work the lowest-lying O(^{1}D) + C_2H_4 PES is further examined at the multiconfigurational self-consistent field (MCSCF), MRMP2, CR-CC(2,3), GVB-PP, and MR-AQCC levels. This study aims to provide a detailed comparison of these different levels of theory for this particular system. In particular, many reactions for this system involve multiple bond rearrangements and require various degrees of both non-dynamic and dynamic correlation for reasonable energetics. As a result of this variety, coupled cluster results parallel but do not always match up with multireference results as previously anticipated. In the case of the CH_2CHOH → oxirane pathway, MCSCF results show the possibility of a two step mechanism rather than an elementary step, but the case is very difficult to elucidate. In the case of the CH_3C:-OH → H_2CCO + H_2 pathway, a non-traditional NEB MEP at the GVB-PP level and MR-AQCC stationary point determination illustrate the need for a complex treatment of this surface.

Introduction:

For the O + ethylene system, both theoretical and experimental[1-5] studies strive towards more accurate potential energy surface (PES) methods, which could subsequently serve as a model to advance combustion studies on larger hydrocarbon systems[6]. Our previous work[1,2] established that discerning transition states (TSs) and minimum energy paths (MEPs) for this system can sometimes be elusive at both the single reference and multireference levels.

Several initial theoretical studies[7-11] of the PES of this reaction were conducted, but the more recent PES studies of Nguyen et al.[12] and Yang et al.[13] gave more comprehensive surveys of the O + ethylene PES. Dynamics studies range from examining the dynamics in the $\bullet CH_2CH_2O\bullet$ biradical region in Hu et al.[14] to single-state dynamics studies of Joshi et al.[15] and Bowman et al.[16-18]

Our first study[1] on the lowest-lying triplet examined pathways 1 – 9 from Nguyen's[12] work with an additional triplet path 10, and our second study[2] analyzed excited states, surface crossings, and reactions near the $\bullet CH_2CH_2O\bullet$ and $\bullet CH_2\bullet CHOH$ biradicals. The current study adds to the previous paper[1] by examining the barriers farther from the $\bullet CH_2CH_2O\bullet$ biradical PES region: the lowest-lying singlet paths 14 – 19 and 21 – 25 as labeled in Nguyen et al.[12] at several levels of theory. This study examines the following reactions at multiple levels of theory.

14) $CH_3CHO \rightarrow CH_2CHOH$

15) $CH_3C{:}\text{-}OH \rightarrow CH_3CHO$

16) $CH_3C{:}\text{-}OH \rightarrow CH_2CHOH$

17) $CH_3CHO \rightarrow \bullet CH_3 + \bullet CHO$

18) $CH_3CHO \rightarrow CH_3{\bullet}CO + {\bullet}H$

19) $CH_3CHO \rightarrow {\bullet}CH_2CHO + {\bullet}H$

21) $CH_3CHO \rightarrow CH_4 + CO$

22) $CH_2CHOH \rightarrow$ oxirane

23) oxirane $\rightarrow CH_3OC{:}\text{-}H$

24) $CH_2CHOH \rightarrow H_2CCO + H_2$

25) $CH_3C{:}\text{-}OH \rightarrow H_2CCO + H_2$

26) oxirane$\rightarrow$ $\bullet\bullet CH_2 + H_2CO$

Methods:

Results from calculations at the MCSCF and MRMP2 levels in this paper were performed with the GAMESS software.[19] The multireference procedure from the previous work[1,2] was used in the current work and is only summarized here. Multireference calculations were carried out at the complete active space self-consistent field (CASSCF)[20-29] level of theory with the aug-cc-pVTZ[30] basis set. The determinant-based method[31] and the full Newton-Raphson converger with augmented Hessian technique[32-34] were used to obtain solutions because of the large complete active space (CAS) sizes. Restricted Hartree-Fock (RHF) or restricted open shell Hartree-Fock (ROHF) calculations with modified valence orbitals[35] and Boys localization[36] provided good starting orbitals for most pathways. In particular, the ROHF orbitals can provide good starting orbitals for open shell singlets. CASSCF stationary point searches for species in the reactions above employed analytical gradients and central-differenced, numerical Hessians. Intrinsic reaction coordinate (IRC) runs were obtained with the second order Gonzalez-Schlegel (GS2) method[37] to identify the minima associated with TSs. If an elementary reaction involves a single bond break with no

TS, a series of constrained optimizations at the CASSCF level was performed to a bimolecular distance of about 5 Å, which was further optimized without constraint as a supermolecule to a bimolecular geometry. The individual pathway sections list any imaginary frequencies present at any bimolecular stationary points with or without a TS in the elementary step.

Once a reaction path was located, second order single state multireference Møller-Plesset perturbation (MRMP2)[38] single point energies using the aug-cc-pVTZ[30] basis set were performed along the entire IRC to recover the majority of the dynamic correlation. From the previous work on the triplet surface, MRMP2 results quantitatively agreed with multireference average quadratic coupled cluster (MR-AQCC)[39] results using COLUMBUS.[40-42] However, previous work[2] on the singlet revealed that this chemical system requires substantial dynamic correlation for some pathways. So, the following additional calculations were performed.

Completely renormalized coupled-cluster singles, doubles, and non-iterative triples CR-CC(2,3)[43,44] single point energies with aug-cc-pVDZ[30] and aug-cc-pVTZ[30] basis sets were also performed for all paths along the IRC (or constrained optimization series) except for 14, 22, and 25. In difficult cases, the ROHF reference was used instead of the RHF reference. For path 14 the GS2 CASSCF IRC calculation failed, and orbital root flipping (ORF– when an orbital not originally in the active space "flips" into the active space) precluded one side of the IRC. So, only stationary points were used in path 14 calculations. Paths 22 and 25 require a more careful examination.

Paths 22 and 25 are examples of particularly difficult portions of the PES. So the generalized valence bond-perfect pairing (GVB-PP)[45-47] approach for the wavefunction was

utilized using COLUMBUS[40-42] for these calculations. For path 25 only, additional MR-AQCC calculations were performed. For the GVB-PP calculations, 9 pairs are used with each pair consisting of two electrons in two orbitals – GVB-PP(9). Besides the chemical core O1s and C1s orbitals, which are doubly occupied, all other orbitals are treated on an equal footing allowing for the correct dissociation of individual electron pairs. The total electronic wave function consists of an antisymmetrized product of the nine perfect pairs yielding a space of 2^9 (512) configurations. At the GVB-PP level, all stationary points of the reaction were characterized by frequency calculations using both the cc-pVDZ[30] and cc-pVTZ[30] basis sets.

On top of the GVB-PP(9)/cc-pVDZ level geometries, a MR-AQCC PP(6) study on path 25 was performed. All pairs are active in the reference space except two non-action CH bonds and the CC σ bond, which are doubly occupied, and their weakly occupied antibonding orbitals in the virtual space. Here and throughout this paper, "action" coordinates refer to internal coordinates that significantly change in length or degree relative to remaining (i.e. "non-action") internal coordinate changes, which are defined by cutoffs. Based on this reference space, single and double excitations are allowed imposing generalized interacting space restrictions,[48] which results in an MR-AQCC expansion space of 6,109,120 configurations. At this level, all stationary points of path 25 were characterized by frequency calculations using the cc-pVDZ[30] basis set. Hessians were computed with the program Suscal[49] by means of finite differences using analytic gradients.[50-52]

The MEP was computed using the nudged elastic band (NEB) method[53-56] for pathway 25 and the traditional IRC method[57-59] for pathway 22 as implemented in MEPPACK (part of COLUMBUS; for more details see Sellner[60]) at the GVB-PP(9)/cc-

pVDZ level. The Hessian matrices used in the relaxation of these NEBs were approximated by scaled diagonal matrices. The convergence threshold on the root-mean-square (RMS) value of the modified gradient on each image in case of the path 25 NEB calculation is 10^{-3} aJ/(Å·sqrt(amu)) and on the RMS value of the projected gradient in each step of the path 22 IRC calculation is 10^{-4} aJ/(Å·sqrt(amu)).

Results and Discussion:

Table 1 displays the active space for each reaction where the labels 'near' and 'far' refer to the carbon closest and farthest from oxygen, respectively. Examination of bond length changes, natural orbital occupation numbers, and localized molecular orbital coefficients over the course of each reaction in different CAS sizes ultimately decide the choice of active space. As described in the previous papers[1-2], results with in-out correlation (IOC – a mostly unoccupied orbital that has an additional, inappropriate radial node) usually result in ORF. So, unless either the IOC constitutes proper radial correlation at some geometry in the pathway or ORF prevents placing the lone O2p in the MCSCF core, CASSCF active spaces do not contain IOC orbitals. However, as will be shown, current, variational techniques do not always allow for chemically intuitive active spaces to persist in many reaction pathways.

Tables 2 and 3 give the overall barrier information for all elementary reactions (except for the multi-step reaction 25) whereby each TS divides each pathway into reactant to TS (forward barrier) and product to TS (reverse barrier). Table 2 shows the CAS barriers with and without zero point energies (ZPE), and Table 3 gives the barriers from MRMP2 energies at the CAS stationary points (stationary point MRMP2 - SPMRMP2), CR-CC(2,3) optimized geometries at the aug-cc-pVDZ[30] basis, Nguyen[12] et al. and Yang[13] et al. barriers,

and single point MRMP2 and CR-CC(2,3) energies at the CASSCF TS geometry and IRC endpoint geometries (i.e. 'select' geometries). While the single point MRMP2 barriers are not strictly barriers of optimized structures, they should be approximately representative of them as previously shown.[1] In addition, for these particular pathways, SPMRMP2 energy barriers always include ZPE (except for species with ORF issues in numerical Hessians as discussed below). Table 4 gives barriers for the GVB-PP and existing MR-AQCC calculations. The natural orbital configuration interaction (NOCI) coefficients help to suggest multireference character. So, in difficult cases, this additional analysis is given. Since no absolute cutoff exists to indicate a multireference situation based on the size of the squared NOCI coefficients, two approximate indicators are chosen for this study. First, if the dominant NOCI weight (i.e. the square of the largest NOCI coefficient) is below 0.95, this approximate cutoff strongly suggests that (dynamical or quasi-degenerate) correlation plays an important role. Paths 14, 21, and 25 each show the dominant NOCI weight below 0.95. Second, if any NOCI weights other than the dominant NOCI weight are above 0.05, this approximate cutoff suggests a multireference situation with quasi-degenerate correlation. Paths 21 and 25 show additional weights above 0.05. Such an analysis might suggest when the CR-CC(2,3) energies are inappropriate. However, even when the above analysis does not suggest multireference character and the reaction of interest simultaneously involves multiple bond rearrangements, borderline cases are possible.

For the rest of the results in this section, specifics on each of the reactions are presented and discussed. Unless otherwise explicitly noted, the descriptions contain the following: barriers and energetics comparisons include ZPE; 'select' barriers are calculated using the select geometries without ZPE; the CASSCF active spaces do not contain IOC and

contain one lone O2p orbital in the MCSCF core; any references made to stationary points refer to CASSCF stationary points (except for explicitly denoted GVB-PP, MR-AQCC, or CR-CC(2,3) optimizations); basis set energy differences are given from the aug-cc-pVDZ[30] to the aug-cc-pVTZ[30] basis at the CR-CC(2,3) level of theory without ZPE at select geometries (as specified in Table 3) and are reported only when the difference exceeds 1.0 kcal/mol; CR-CC(2,3) select barriers are reported with the aug-cc-pVTZ basis in the text; and individual pathways report only non-negligible geometry changes in this study's CASSCF stationary points relative to Nguyen's[12] B3LYP results.

Pathway 14: $CH_3CHO \rightarrow CH_2CHOH$

In this reaction a hydrogen shifts from the carbon farthest from the oxygen to the oxygen itself. The involvement of the oxygen here in multiple bond changes (e.g. the transformation of a CO π, π^* to a CC π, π^*) requires three O2p orbitals in the active space and leads to IOC, which requires the use of an even (14,14) CAS. Both minima are A′ states with C_s symmetry while the TS is asymmetric. Despite the presence of both a reasonable active space and frequencies at the TS, the GS2 computation resulted in an excess of constrained optimizations on the hypersphere surface in an attempt to reach the next point in the IRC. In addition, the vinyl alcohol ends up with a carbon-carbon IOC orbital instead of an O2p* IOC orbital in the CAS. Therefore, an IRC is unavailable for this pathway.

The CR-CC(2,3) TS optimization failed while CR-CC(2,3) select barriers do not even match any other results to within ~10.0 kcal/mol. Because of this failure, further analysis is required in terms of weights (as previously described). At the CASSCF TS geometry, the dominant NOCI weight is 0.90 with no other weights above the approximate 0.05 cutoff, and the coupled cluster T2 amplitude magnitudes are small (i.e. | amplitude | < 0.1 as the cutoff).

However, as stated above, this reaction simultaneously involves multiple bond changes which suggests that multireference character may be playing a role. The forward (reverse) barrier is 68.1 (57.1) kcal/mol at the SPMRMP2 level. Because of the carbon-carbon IOC orbital on vinyl alcohol, the reverse barrier could be less accurate than the forward barrier. In addition, vinyl alcohol has two possible conformations with C_s symmetry because of OH rotation, which gives an energy difference of only ~1 kcal/mol at the CASSCF level. Since an IRC for a different reaction with a smaller CAS space leads to the higher energy conformer, the enthalpy was calculated with the higher energy conformer. Experimental results show the overall reaction enthalpy to be 9.9 kcal/mol,[61,62] which closely matches 11.0 (11.3) kcal/mol derived from SPMRMP2 (CR-CC(2,3) optimization) for this path. Furthermore, these results also match Nguyen and Yang results to within 1.0 kcal/mol.

Pathway 15: CH_3C:-OH → CH_3CHO

For this pathway an (8,8) CAS models a hydrogen shift from 1-hydroxyethylidene to acetaldehyde. In this case, the IRC has C_s symmetry. Despite several attempts at a (6,6) and an (8,7) CAS, this pathway unfortunately required an IOC orbital in order to avoid ORF. Dynamic correlation from MRMP2 leads to a modest shift in pathway energetics and essentially no shift along the reaction coordinate. Here and throughout this paper, a shift simply refers to how single point energies at a different level of theory horizontally and vertically change the IRC curvature. The forward barrier (reverse barrier) is 23.3 (78.2) kcal/mol at the SPMRMP2 level, and with CR-CC(2,3) optimized geometries, the forward (reverse) barrier is 29.3 (78.6) kcal/mol. Nguyen, Yang, and the CR-CC(2,3) barriers all match to within 1 kcal/mol.

Pathway 16: CH_3C:-OH → CH_2CHOH

The reaction for 1-hydroxyethylidene to vinyl alcohol requires at least a (10,10) CAS. In addition, a (14,14) CAS provides some additional comparisons for this pathway. The (14,14) CAS has IOC whereas the (10,10) CAS has a core O2p (i.e. no IOC). Here, both minima have C_s symmetry while the rest of the pathway has no symmetry. For the (10,10) CAS with the lower energy vinyl alcohol conformer as an end point, the forward barrier (reverse barrier) is 19.0 (62.7) kcal/mol at the SPMRMP2 level. For the reverse barrier in the (14,14) CAS, the IRC converges to the vinyl alcohol minimum with a rotated hydroxyl group in a higher energy conformation; but just as in path 14, the rotation makes little difference to the reverse barrier. Based on the CR-CC(2,3) data, the basis set energy difference is approximately 1.1 kcal/mol. The CR-CC(2,3) TS optimization failed. The T2 amplitude magnitudes are below the cutoff, and the dominant NOCI weight is 0.93. The CR-CC(2,3) select forward (reverse) barrier is 25.0 (64.4) kcal/mol. Here, the select barriers involve the higher energy vinyl alcohol conformer. These results match Nguyen and Yang forward barrier to within 2 kcal/mol.

For geometry comparisons, (10,10) and (14,14) CAS geometries are very similar. The CC bond length goes from 1.53 to 1.41 Å in the forward barrier and from 1.41 to 1.35 Å in the reverse barrier. The TS geometry has action CCH angles of 60 ^{o} and 56 ^{o}. In the forward barrier, the CO distance slightly changes from 1.30 to 1.35 Å.

<u>Pathway 17:</u> $CH_3CHO \rightarrow \bullet CH_3 + \bullet CHO$

This pathway is the first of three barrierless, acetaldehyde dissociations that require both constrained optimizations and a basis larger than a double zeta basis. Here, barrierless means that there is no maximum in the PES before dissociation. For this pathway an (8,8) CAS provides a qualitatively reasonable wavefunction for breaking the carbon-carbon bond

plus accompanying bonding changes. From previous calculations with the analogous triplet pathway,[1] it is well known that the active space must contain all three O2p orbitals. Since ORF occurs for a pure (i.e. not state-averaged) state (8,7) CAS, the calculations must include IOC. The results for this pathway are all in the A′ state in C_s symmetry. Since this pathway has no TS, energy differences between minima are given.

First, based on the CR-CC(2,3) select barriers, this pathway displays a 4.2 kcal/mol basis set energy difference. So, a double zeta basis would be insufficient to describe the energetics to within 1 kcal/mol. Second, at the SPMRMP2 level, this separation requires 74.6 kcal/mol. The 5 Å fragment (not a CASSCF minimum within the constraints of the constrained optimization) has one imaginary frequency of 26*i* cm^{-1} at the CASSCF level that points along the motion of the two products leaving each other (along the line of symmetry). Extending the distance to ~10 Å, the supermolecule energy only differs by ~1 kcal/mol from the 5 Å supermolecule energy. Third, the CR-CC(2,3)/aug-cc-pVDZ level, supermolecule geometry optimizations led to an energy difference of 84.4 kcal/mol at 5 Å distance while optimizations on the individual doublet CH_3 and CHO products and singlet acetaldehyde yielded differences of 85.8 kcal/mol (without ZPE) and 77.9 kcal/mol (with ZPE). More strictly comparing select MRMP2 and CR-CC(2,3) energies confirms the particular importance of dynamic correlation for accurate energetics in this pathway. Comparing the supermolecule and doublet optimized energetics shows the large contribution to the dynamic correlation derives from interactions from the CC σ, σ* bonds. In fact, at the 5 Å supermolecule, T2 amplitude magnitudes at or above the 0.1 cutoff are on the orders of 0.75 and 0.1, which indicate strong multireference character.

The 5 Å CR-CC(2,3), 84.4 kcal/mol energy difference matches the experimental result of 84.8 kcal/mol[63] and Nguyen energetics of 82.0 kcal/mol. With respect to the SPMRMP2 energetics, IOC in the wavefunction might slightly contribute to the 10.2 kcal/mol error in the final energetics. In terms of energy differences with respect to differing geometries, the individual coupled cluster geometries differ by very little from the corresponding geometries in a 5 Å supermolecule. Given the difference in the basis set energies and the computationally expensive CR-CC(2,3) optimizations with a triple zeta basis, these coupled cluster calculations give the barrier to be ~75.2 – 80.6 kcal/mol, which matches the MRMP2 results just as well as the experimental barrier. It should also be noted that while there are large energy barriers in these reactions, they will be favored by entropy since they lead to multiple products.

Pathway 18: $CH_3CHO \rightarrow CH_3{\bullet}CO + {\bullet}H$

This hydrogen dissociation also requires constrained optimizations in C_s symmetry with an (8,8) CAS that contains IOC; modeling this pathway involves the same issues as modeling pathway 17. The basis set energy difference is 3.4 kcal/mol, which is again quite large. At the ~5 Å product, two imaginary frequencies are still present for the CASSCF, constrained stationary point. The first imaginary frequency of $34i$ cm^{-1} points from the hydrogen fragment along the motion of the two fragments leaving each other (along the line of symmetry). The second imaginary frequency of $33i$ cm^{-1} involves a methyl rotation and a bend of the CO group out of the plane. In addition, the CR-CC(2,3)/aug-cc-pVDZ level, supermolecule geometry optimizations (again, with two small imaginary frequencies) led to an energy difference of 86.8 kcal/mol, and the CR-CC(2,3)/aug-cc-pVDZ level, doublet

optimizations led to an energy difference of 91.2 kcal/mol (without ZPE) and 83.6 kcal/mol (with ZPE).

As a result, the separation requires 83.5 kcal/mol at the SPMRMP2 level while the CR-C(2,3)/aug-cc-pVDZ supermolecule energy difference of 86.8 kcal/mol more closely matches Nguyen energy difference of 87.3 kcal/mol and the experimental barrier of 89.3 kcal/mol.[61,64,65] Again, these MRMP2 calculations did not recover enough dynamic correlation to get the energetics correct for this bimolecular reaction while the CR-CC(2,3) calculations will have issues due to the multireference character of the wavefunction. Besides the dissociating hydrogen, no major geometry changes take place in this pathway.

Pathway 19: $CH_3CHO \rightarrow \bullet CH_2CHO + \bullet H$

This hydrogen dissociation is the last barrierless pathway that was calculated with constrained optimizations. With a basis set energy difference of 3.9 kcal/mol, a larger than double zeta basis is again required for this pathway. Our previous paper[1] gives this dissociation on the triplet surface. The triplet dissociation is extremely problematic both with and without IOC; only an active space of three O2p orbitals and no IOC could lead to a smooth IRC. However, as before, (N,N-1) pure state, non-full valence, active spaces are not feasible with current convergers and still lead to ORF. Thus, for the singlet dissociation, the MCSCF core contains a lone O2p with the understanding that this CAS will lead to the least convergence issues in the search for pure states in this constrained optimization series. This pathway requires at least a (12,12) CAS and is asymmetric except at acetaldehyde, for which the calculations do not use symmetry. Pulling the fragments out to ~5 Å yields a separation energy of 88.6 kcal/mol at the SPMRMP2 level. CC-CR(2,3) optimized geometries result in energy differences of 90.3 kcal/mol for fully separated doublet products. CR-CC(2,3)

supermolecular optimizations yielded diverging energies and were not pursued further. These results differ from Nguyen's energetics of 93.8 kcal/mol by 3.5 kcal/mol. In fact, relative to the Nguyen values, the CR-CC(2,3) select energy difference is larger at 103.8 kcal/mol.

Pathway 21: $CH_3CHO \rightarrow CH_4 + :CO$

This pathway from acetaldehyde also requires a compromised CAS because of computational expense. A (10,10) CAS with IOC starts with and retains acetaldehyde's methyl group in the MCSCF core rather than an active space that additionally contains the methyl group (i.e. compromised CAS). The reaction ultimately results in methane and carbon monoxide bimolecular fragments. For this pathway both acetaldehyde and ~5 Å bimolecular fragments calculations have an A′ state in C_s symmetry; so this pathway's IRC appears to be entirely symmetric with the given CAS size. However, the symmetric fragment geometry contains a small imaginary frequency of $22i$ cm^{-1} that essentially rotates the entire methane. Without this symmetry restriction on the bimolecular fragments, an ORF occurs so that all methane CH σ orbitals equivalently reside in the core. The forward barrier (reverse barrier) is 76.9 (84.7) kcal/mol at the SPMRMP2 level. CR-CC(2,3) optimized geometries result in a forward (reverse) barrier of 82.5 (90.8) kcal/mol. However, the dominant CASSCF TS NOCI weight is 0.87 while another weight is 0.07, which indicates the correlation is important. In addition, the occupation number also suggests multireference character with a value of 0.17. Furthermore, the CR-CC(2,3) optimized doublets demonstrate that further separation comprises less than 1 kcal/mol contribution to the reverse barrier. The reaction enthalpy of -7.8 (-8.3) kcal/mol at the SPMRMP2 (CR-CC(2,3)/aug-cc-p-VDZ supermolecule) level differs from the experimental reaction enthalpy of -14.5

kcal/mol.[61,66] Because of the CR-CC(2,3) results, this difference from experiment most likely does not derive from the use of the compromised CAS. Nguyen's forward and reverse barriers both lie within 2 kcal/mol of the corresponding CR-CC(2,3) results.

The OCC angle changes from 125 ^{o} to 107 ^{o} in the course of the forward barrier. The CC σ bond breaks in the forward barrier; and the CC distance changes from 1.53 to 2.22 Å. For the TS this study gives the distance from the action H to the carbon of methyl as 1.89 Å while Nguyen's study gives 1.73 Å; as well, this study finds the TS CC distance as 2.22 Å while Nguyen shows it to be 2.10 Å.

Pathway 22: $CH_2CHOH \rightarrow$ oxirane

Nguyen et al. reported this rearrangement as a concerted reaction. However, Yang et al.[13] indicated that at the CCSD(T)/cc-pVTZ//B3LYP/6-311++G* level, this pathway does not exist and in fact requires two steps: $CH_2CHOH \rightarrow HC{:}\text{-}CH_2OH$ and $HC{:}\text{-}CH_2OH \rightarrow$ oxirane. This reaction involves multiple bond formations and multiple numbers of lone pairs (i.e. O2p, C2p). In addition, the CR-CC(2,3) TS optimization also failed, which again suggests difficulties with this level of theory for multiple bond rearrangements. All attempts to find a one step mechanism failed. However, in order to obtain energetics for the desired intermediates for the total reaction, GVB-PP geometry optimizations started with Nguyen's reported stationary points and produced corresponding extrema. GVB-PP gives the overall reaction with an enthalpy of 18.2 kcal/mol, which matches 17.2 kcal/mol from experiment.[62,66] In an attempt to clarify the elementary steps for this reaction, various CAS sizes were used, but this study only reports results for non-ORF CAS and GVB-PP results below.

Pathway 22 Step 1: $CH_2CHOH \rightarrow HC{:}\text{-}CH_2OH$

Locating an IRC at the density function theory with the hybrid B3LYP[67,68] functional and the 6-31G*[69,70] basis provided reasonable initial geometries for CASSCF geometry searches. Then, plotting CR-CC(2,3) and (6,6)CAS single point energies along the DFT IRC reveals a particular difference in the elongated curvature (see Figure 1): the larger barrier does not exist in the CASSCF curve unlike the single reference DFT and CR-CC(2,3).

In order to properly model this hydrogen transfer, a small, compromised (6,6)CAS size is required. Even though a possibly more chemically intuitive, larger CAS size is computationally feasible in this case, ORF occurs during the course of the IRC. At the beginning of this IRC, the active space minimally includes the CC and CH σ spaces and the CC π space. Figure 2 shows the CASSCF IRC with the MRMP2 and CR-CC(2,3) single point energies. The HC:-CH_2OH endpoint of the CASSCF IRC roughly corresponds to the end of the first barrier shown in the DFT IRC (Figure 1) (around s=-2 bohr·sqrt(amu)).

Based on the CR-CC(2,3) data, the basis set energy difference is approximately 1.9 kcal/mol. The GVB-PP(9)/cc-pVDZ IRC forward (reverse) barrier is 74.4 (7.5) kcal/mol, the (6,6)CAS IRC forward (reverse) barrier is 79.1 (0.7) kcal/mol, the forward (reverse) barrier is 77.7 (-4.1) kcal/mol at the SPMRMP2 level, and the CR-CC(2,3) select forward (reverse) barrier is 76.4 (-3.1) kcal/mol. Because of the curvature differences between DFT, CR-CC(2,3), and (6,6)CAS in Figure 1, CR-CC(2,3) TS optimizations were not attempted. It is important to note that both the SPMRMP2 and CR-CC(2,3) results suggest that adding dynamic correlation makes this barrier disappear.

Pathway 22 Step 2: HC:-CH_2OH → oxirane

For the second step, a compromised (10,10)CAS is used for modeling. In particular, on the reactant side of step 2, the 3 non-action CH σ orbitals are left out of the active space.

Figure 3 shows the CASSCF IRC with the MRMP2 and CR-CC(2,3) single point energies. Both the MRMP2 and CR-CC(2,3) methods lower the barriers relative to the CASSCF method. The HC:-CH_2OH endpoint of the CASSCF IRC is not the same as the one for step 2 and involves rotation of the O-C-C-H and C-C-O-H dihedrals. By performing a scan between these two points at the CASSCF (SPMRMP2) level, the step 2 endpoint is ~6.6 (7.1) kcal/mol uphill in energy from the step 1 endpoint without an explicit barrier.

Based on the CR-CC(2,3) data, the basis set energy difference is again approximately 1.9 kcal/mol. The GVB-PP(9)/cc-pVDZ IRC forward (reverse) barrier is 28.6 (78.8) kcal/mol, the (10,10)CAS forward (reverse) barrier is 10.7 (75.4) kcal/mol, the forward (reverse) barrier is -1.3 (68.0) kcal/mol at the SPMRMP2 level, and the CR-CC(2,3) select forward (reverse) barrier is 3.4 (72.2) kcal/mol. The large GVB-PP forward barrier relative to the CAS and SPMRMP2 barriers most likely occurs since GVB-PP does not include the majority of the quasi-degenerate correlation for this step, which involves action among several more bonds than even in the reverse barrier for path 22 step 1. So, while the CASSCF level shows two barriers (Figures 2 and 3), the CR-CC(2,3) and SPMRMP2 data suggest that that the overall reaction (i.e. pathway 22) is one high energy barrier with a very flat, extended surface (Figures 2 and 3 combined) that may or may not involve two steps.

Pathway 23: oxirane → CH_3OC:-H

This pathway depicts the ring opening of oxirane and a hydrogen shift as concerted processes. As such, this compromised (10,10) CAS keeps 3 CH σ orbitals in the core and has IOC; this active space choice does not allow C_s symmetry for oxirane. As well, symmetric A′ calculations on CH_3OC:-H were not attempted because of previous difficulties with compromised CAS starting orbitals. Various attempts to place an O2p into the core (and

hence allow for more CH σ, σ* oribtals in the CAS) fail because the O2p enters the active space once the IRC reaches the CH_3OC:-H intermediate. Although this pathway has two closed shell endpoints, it differs from other 'closed shell' pathways in the fact that dynamic correlation slightly shifts the TS along the reaction coordinate; the minima only energetically shift. The forward barrier (reverse barrier) is 74.9 (27.1) kcal/mol at the SPMRMP2 level. As well, this pathway shows a small, 1.5 kcal/mol basis set energy difference. Even though the NOCI and T2 cutoffs do not indicate a multireference situation, this reaction involves multiple bond changes, and the dominant NOCI weight is 0.92 for the TS.

The CR-CC(2,3) TS optimization also failed for this reaction. So, the CR-CC(2,3) select forward (reverse) barrier is 77.9 (32.6) kcal/mol. SPMRMP2 forward (reverse) barrier lies 1.0 (8.8) kcal/mol shallower than the corresponding Nguyen barrier while CR-CC(2,3)/aug-cc-pVTZ select forward (reverse) barrier lies 2.1 kcal/mol deeper (3.3 kcal/mol shallower) than the corresponding Nguyen barrier.

Pathway 24: $CH_2CHOH \rightarrow H_2CCO + H_2$

For this hydrogen dissociation from vinyl alcohol, a (10,10) CAS has two CH σ orbitals farthest from the oxygen in the core as a compromise and also has IOC. The reaction stationary points all have C_s symmetry with an A′ state. However, the CCO angle in H_2CCO slightly bends at the ~4 Å product end. From our previous study,[2] pathway 20 also has this common minimum but does not suffer from this bent angle with a (12,12) CAS. Furthermore, this part of the surface tends to be very flat, which makes optimization difficult. Thus, this bimolecular ~4 Å geometry was pulled out to ~10 Å and asymmetrically re-optimized to avoid this issue. Comparing the ~10 Å and the geometry at the end of the IRC gives less than 1 kcal/mol change in the CASSCF energy. In this case, MRMP2 gives a

horizontal shift for the bimolecular ~4 Å product. The forward barrier (reverse barrier) is 84.4 (68.3) kcal/mol at the SPMRMP2 level. Furthermore, the CR-CC(2,3) optimizations led to a forward (reverse) barrier of 85.6 (68.2) kcal/mol. This forward barrier matches the Nguyen forward barrier within 1.0 kcal/mol and the SPMRMP2 barrier within 1.2 kcal/mol. The reverse barrier matches the SPMRMP2 barrier within 1.0 kcal/mol but lies 2.4 kcal/mol shallower than the Nguyen reverse barrier.

Pathway 25: CH_3C:-OH $\rightarrow$ $H_2CCO + H_2$

This hydrogen molecular dissociation presents a challenge to multireference calculations. After initial calculations with various CAS sizes from Nguyen's starting saddle point geometry, a (12,12) CAS provides a compromised solution where two out of the three, degenerate CH σ orbitals at the unimolecular intermediate remain in the MCSCF core throughout the saddle point search. However, such a search resulted in an unexpected saddle point geometry, or TS 25 (Figure 4). In addition, the CR-CC(2,3) TS optimization failed, and the dominant NOCI weight is 0.84. Another NOCI weight is 0.08, and an occupation number of 0.21 is found, which possibly indicates some multireference character. In order to further characterize this PES region, GVB-PP calculations with the NEB method provided a non-traditional MEP that involves several TSs. Figure 4 displays the NEB results; the potential near TS 25 is particularly flat at the GVB-PP level of theory. Here, the non-traditional MEP between the reactants and products involves three saddle points, and the surface flatness makes location of the minima between these saddle points very difficult. In order to determine the effect of dynamic correlation on this path, stationary points were located at the MR-AQCC level of theory with a cc-pVDZ basis. As a result of the flat surface, TS 25 strongly depends on the presence or absence of dynamic correlation. In

particular, the OH distance is 3.76 Å at the GVB-PP level whereas the same OH distance is 1.95 Å at the MR-AQCC level. Table 4 gives the forward barrier to TS OH at the GVB-PP level and TS 25 at the MR-AQCC level. The forward barrier is 46.8 kcal/mol and 33.0 kcal/mol at the GVB-PP and MR-AQCC levels, respectively.

The most noticeable geometry difference between these structures and Nguyen's saddle point geometry lies in the HH distance. The TS OH geometry most resembles Nguyen's saddle point; however, GVB-PP (MR-AQCC) level of theory yields an HH distance of 2.16 (2.01) Å whereas Nguyen reported 1.08 Å. As well, the MR-AQCC TS 25 geometry gives a more comparable HH distance of 1.15 Å, but as already noted above, the OH distance differs from Nguyen's reported geometry. Overall, dynamic correlation is as important as quasi-degenerate correlation for determining an accurate, TS geometry in this case.

Pathway 26: oxirane→ $\bullet\bullet CH_2 + H_2CO$

For this pathway a (6,6) CAS provides the IRC, which has C_s symmetry throughout the forward barrier and most of the reverse barrier. For the energetics with respect to the basis set, this pathway shows a large 4.6 kcal/mol basis set energy difference. Including dynamic correlation at the SPMRMP2 level leads to a barrierless reaction with an energy gain of 81.4 kcal/mol. Including more correlation with CC-CR(2,3)/aug-cc-pVDZ doublet optimizations resulted in a energy difference of 79.6 kcal/mol. Based on the above basis set energy difference, the energy difference could approximately vary towards 84.2 kcal/mol, which is similar to the Nguyen result of 85.7 kcal/mol. However, SPMRMP2/aug-cc-pVTZ single point energy results agree more with the CR-CC(2,3)/aug-cc-pVDZ optimization results than the Nguyen results. No major differences exist in the fragmented geometries.

The CASSCF has a potential well on the product side at ~3 Å separation (reaching this symmetric product requires a rotation of the formaldehyde with respect to the methylene, which breaks the original pathway symmetry). The SPMRMP2 reaction enthalpy of 81.4 kcal/mol is within 1.3 kcal/mol of the experimental enthalpy of 80.1 kcal/mol.[71-73]

Conclusions:

In this work, selected pathways of the $O(^1D) + C_2H_4$ PES were calculated at the CASSCF, MRMP2, and CR-CC(2,3) levels of theory. Whenever possible, CASSCF active spaces for pure states did not contain lone O2p orbitals, which tend to lead to IOC ORF with the current, convergence techniques. This active space construction is particularly important in both solving for pure states at the CASSCF level (avoiding IOC ORF) and obtaining single state MRMP2 results (avoiding state root flipping). In addition, dynamic correlation is also essential for obtaining reasonable energetics – horizontal shifts in the locations of TSs and minima using single point energetics are small for the singlet reactions in this study. The MRMP2 treatment captures a significant portion of this correlation but certainly does not encapsulate all of it. In most cases considered here, CR-CC(2,3) results recover additional correlation missed by MRMP2, but in some cases, these coupled cluster results appear to be inappropriate as indicated by the NOCI coefficients. Nonetheless, given comparisons of CR-CC(2,3) single point data to both Nguyen and experiment, CR-CC(2,3) optimized barriers appear to be crucial for accurate energetics. Furthermore, CR-CC(2,3) TS optimizations sometimes fail for multiple bond changes in some reactions, for which NOCI expansions sometimes suggest multireference character. For a variety of reasons, some possibly difficult cases are paths 14, 16, 21, 23, and 25. These cases and the cases with a possibly insufficient, double zeta basis in the CR-CC(2,3) calculations limit valid conclusions.

Comparing this work's enthalpies with those available from experiment gives reasonable agreement in many cases. In these reactions, the calculated enthalpy errors range from less than 2.5 kcal/mol to a single, worst case of 6.2 kcal/mol. However, in one case (i.e. pathway 21), our results compare well with Nguyen et al. Future work on O + ethylene will take a number of directions that include further GVB-PP and MR-AQCC energetics with eventual applications towards dynamics for this system.

Acknowledgments:

The authors are indebted to Michael W. Schmidt and Mark S. Gordon for help in using the capabilities of GAMESS and MCSCF. This material is based upon work supported by the National Science Foundation under Grant No. OISE-0730114 for the Partnerships in International Research and Education (PIRE) and by the Robert A. Welch Foundation under Grant No. D-0005. TeraGrid resources were provided by the Texas Advanced Computing Center (TACC). Support was also provided by the High-Performance Computing Center (HPCC) at Texas Tech University, under the direction of Philip W. Smith. In addition, this work was supported by the Austrian Science Fund within the framework of the Special Research Program F41 (Vienna Computational Materials Laboratory (ViCoM)). Computer time at the Vienna Scientific Cluster (project no. 70019) is gratefully acknowledged. TLW acknowledges computing resources purchased through funds provided by Ames Laboratory and Iowa State University.

Electronic Supplemental Material:

The online version of this article contains supplementary material, which is available to authorized users.

Table 1. Description of the main CAS sizes used for each reaction pathway. Only the lone pairs and bonding orbitals are shown; the antibonding orbitals can be inferred from bonding orbitals and the indication of in-out correlation. The O2s is always in the core.

reaction pathway	active space	in-out correlation	description
14	(14,14)	Y	3 CHσ, COπ, COσ, CCσ, lone O2p →2 CHσ, CCπ, OHσ, COσ, CCσ, lone O2p
15	(8,8)	Y	COσ, COπ, OHσ, lone C2p → COσ, COπ, CHσ, lone O2p
16	(10,10)	N	CCσ, 3 CHσ, lone C2p → CCσ, 3 CHσ, CCπ
	(14,14)	Y	COσ, CCσ, 3 CHσ, lone C2p, lone O2p → COσ, CCσ, 3 CHσ, CCπ, lone O2p
17	(8,8)	Y	COπ, lone O2p, COσ, CCσ → 2 COπ, COσ, biradical(near C2p;O2p, far C2p)
18	(8,8)	Y	COπ, lone O2p, COσ, CHσ → 2 COπ, COσ, biradical(near C2p;O2p, H)
19	(12,12)	N	3 CHσ, COπ, COσ, CCσ → 2 CHσ, COπ, COσ, CCσ, biradical(far C2p, H)
21	(10,10)	Y	near CHσ, COπ, loneO2p, COσ, CCσ → far CHσ, COπ, COπ, COσ, CCσ

reaction pathway	active space	in-out correlation	description
22 Step 1	(6,6)	N	far CHσ, CCσ, CCπ → near CHσ, CCσ, lone C2p, lone C2s
22 Step 2	(10,10)	Y	OHσ, CCσ, COσ, lone C2p,lone O2p → CHσ, CCσ, 2COσ, lone O2p
23	(10,10)	Y	CHσ, CCσ, 2COσ, lone O2p → CHσ, lone C2p, 2COσ, COπ
24	(10,10)	Y	CHσ, OHσ, loneO2p, CCπ, COσ → HHσ, 2 COπ, CCπ, COσ
26	(6,6)	N	CCσ, 2COσ → COσ, COπ, biradical(far C2p, far C2s)

Table 2. Barriers to reactions in kcal/mol calculated at the CAS level without and with (in parenthesis) ZPE. Nguyen's barriers include ZPE.

reaction pathway	CAS forward barrier[a]	CAS reverse barrier[a]	Nguyen forward barrier[b]	Nguyen reverse barrier[b]
14	80.1 (76.8)	51.2 (47.4)	67.6	56.9
15	34.2 (30.4)	90.9 (87.2)	28.4	79.4
16[c]	27.0 (24.6)	72.0 (69.1)	23.2	63.5
16[d]	29.6 (27.0)	70.1 (67.2)		
17	76.7 (68.9)	N/A (N/A)	82.0	N/A
18	85.5 (78.0)	N/A (N/A)	87.3	N/A
19	92.7 (84.6)	N/A (N/A)	93.8	N/A
21	88.1 (83.6)	102.5 (101.7)	83.4	89.3
22 Step 1	81.7 (79.1)	1.4 (0.7)	N/A	N/A
22 Step 2	12.4 (10.7)	79.0 (75.4)	N/A	N/A
23	82.2 (78.5)	39.0 (37.2)	75.8	35.9
24	100.3 (94.3)	83.4 (86.8)	85.6	70.6
26	3.6 (4.6)	70.1 (65.0)	N/A	N/A

a. CAS barriers are obtained from unconstrained optimizations of intermediates and ~3-5 Å separated products.

b. Nguyen's[12] barriers are the average barriers from several different levels of theory.

c. (10,10) CAS.

d. (14,14) CAS.

Table 3. Barriers to reactions in kcal/mol from single point energies calculated at the MRMP2 level with aug-cc-pVTZ, and CR-CC(2,3) level with both aug-cc-pVDZ (ACCD) and aug-cc-pVTZ. F indicates the forward barrier while R indicates the reverse barrier. SPMRMP2 barriers include ZPE from CASSCF stationary points. Nguyen's barriers include ZPE. Select MRMP2 and CR-CC(2,3) barriers do not include ZPE and are all at the same CASSCF geometries. In addition, energy differences without and with CR-CC(2,3) ZPE are given for geometry optimizations at the CR-CC(2,3) level with an aug-cc-pVDZ basis.

reaction pathway	SPMRMP2[a] (ZPE)	Opt CR-CC(2,3) (ZPE) (ACCD)	Nguyen[b] /Yang (ZPE)	Opt CR-CC(2,3) (ACCD)	Select MRMP2	Select CR-CC(2,3)	Select CR-CC(2,3) (ACCD)
14 R	57.1	-	56.9 56.4[d]	-	60.9[f]	70.5[f]	71.0[f]
14 F	68.1	-	67.6 67.0[d]	-	71.4[f]	59.6[f]	59.0[f]
15 R	78.2	78.6	79.4 79.3[d]	82.6	81.9	82.5	82.7
15 F	23.3	29.3	28.4 28.7[d]	33.0	27.0	32.3	32.9
16 R[i]	62.7	-	63.5 62.4[d]	-	64.7	64.4	63.3
16 F[i]	19.0	-	23.2 23.7[d]	-	21.4	25.0	25.4
16 R[j]	61.2	-	-	-	64.1	-	-
16 F[j]	19.5	-	-	-	22.1	-	-
17	74.6	84.4 77.9[g]	82.0	92.0 85.8[g]	82.4	96.6	92.4

reaction pathway	SPMRMP2[a] (ZPE)	Opt CR-CC(2,3) (ZPE) (ACCD)	Nguyen[b] /Yang (ZPE)	Opt CR-CC(2,3) (ACCD)	Select MRMP2	Select CR-CC(2,3)	Select CR-CC(2,3) (ACCD)
18	83.5	86.8 83.6[g,h]	87.3	94.3 91.2[g,h]	91.0	97.8	94.4
19	88.6	- 90.3[g,h]	93.8	- 98.3[g,h]	96.7	103.8	99.9
21 R	84.7	90.8 91.1[g]	89.3	91.9 91.7[g]	85.6	91.6	92.5
21 F	76.9	82.5	83.4	86.8	80.1	86.9	86.5
22 S1 R[k]	-4.1	-	-	-	-3.4	-3.1	-2.6
22 S1 F[k]	77.7	-	-	-	80.3	76.4	74.5
22 S2 R[k]	68.0	-	-	-	71.6	72.2	70.3
22 S2 F[k]	-1.3	-	-	-	0.4	3.4	5.0
23 R	27.1 29.0[c]	-	35.9 32.1[d]	-	28.9	32.6	33.3
23 F	74.9 76.8[c]	-	75.8 74.9[d]	-	78.5	77.9	76.4
24 R	68.3	68.2	70.6	66.0	65.3	65.9	65.7
24 F	84.4	85.6	85.6	91.8	89.6	90.3	90.6
26	81.4	- 79.6[g]	85.7	- 88.5[g]	91.6	89.9	85.3

a. SPMRMP2 values derive from CAS stationary points, which are obtained from unconstrained optimizations of intermediates and ~3-5 Å separated products.

b. Nguyen's[12] barriers are the average barriers from several different levels of theory and are unlabeled in the column.

c. MRMP2 barrier with ZPE from shifts in the TS only.

d. Yang's[13] barriers at the CCSD(T)/cc-pVTZ//B3LYP/6-311++G* level of theory.

e. ORF incidates orbital root flipping.

f. Single point energies performed at CASSCF stationary points rather than off of IRC.

g. Optimizations separately performed on each product fragment.

h. Energy value for doublet hydrogen is taken from ROHF as -0.4993343154 Hartree.

i. (10,10) CAS.

j. (14,14) CAS.

k. "S" indicates step for the two pseudo-elementary steps located for pathway 22.

Table 4. Barriers to reactions in kcal/mol calculated at the GVB-PP level without and with (in parenthesis) ZPE at the cc-pVDZ or cc-pVTZ basis as labeled. Path 25 also includes barriers with MR-AQCC optimized geometries at the cc-pVDZ basis with ZPE in square brackets; as described in the text, the TS geometry with the most reasonable geometry is used in the construction of path 25 barriers. Nguyen's barriers include ZPE.

reaction pathway	GVB-PP forward barrier[a]	GVB-PP reverse barrier[a]	Nguyen forward barrier[b]	Nguyen reverse barrier[b]
22 S1[d,e]	77.0 (74.4)	8.0 (7.5)	N/A	N/A
22 S2[d,e]	29.4 (28.6)	82.5 (78.8)	N/A	N/A
25	51.8[f] (46.6)[f] [33.0][c,e]	N/A (N/A) [N/A]	33.0	58.3

a. GVB-PP barriers from GVB-PP stationary points.

b. Nguyen's[12] barriers are the average barriers from several different levels of theory.

c. MR-AQCC barrier from MR-AQCC stationary points.

d. "S" indicates step for the two pseudo-elementary steps located for pathway 22.

e. cc-pVDZ basis

f. cc-pVTZ basis

Figure 1. Path 22 Step 1: DFT IRC and Single Point Energies (6-31G* basis no ZPE)

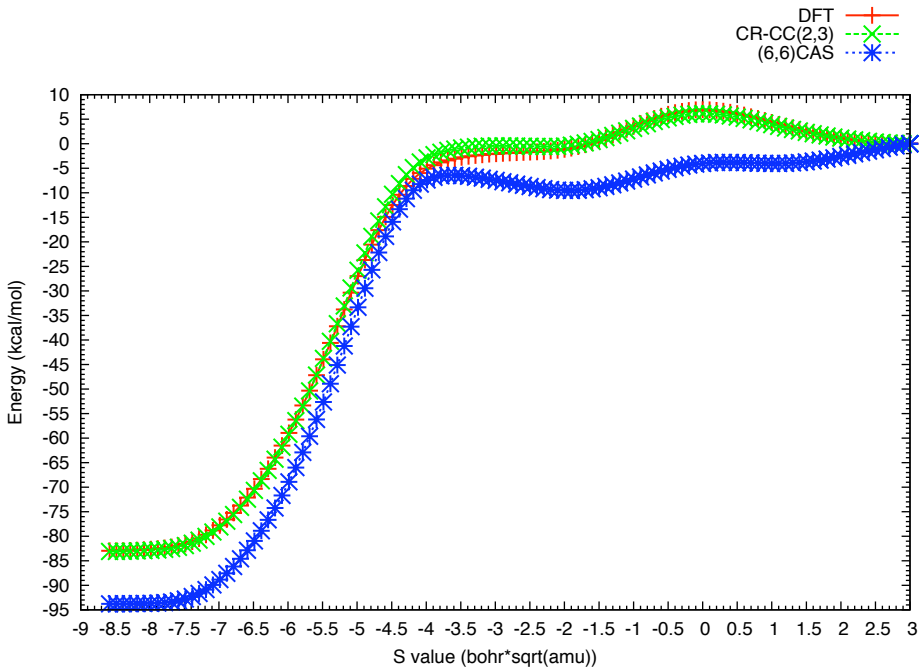


Figure 2. Path 22 Step 1: CASSCF IRC and Single Point Energies (aug-cc-pVTZ basis no ZPE)

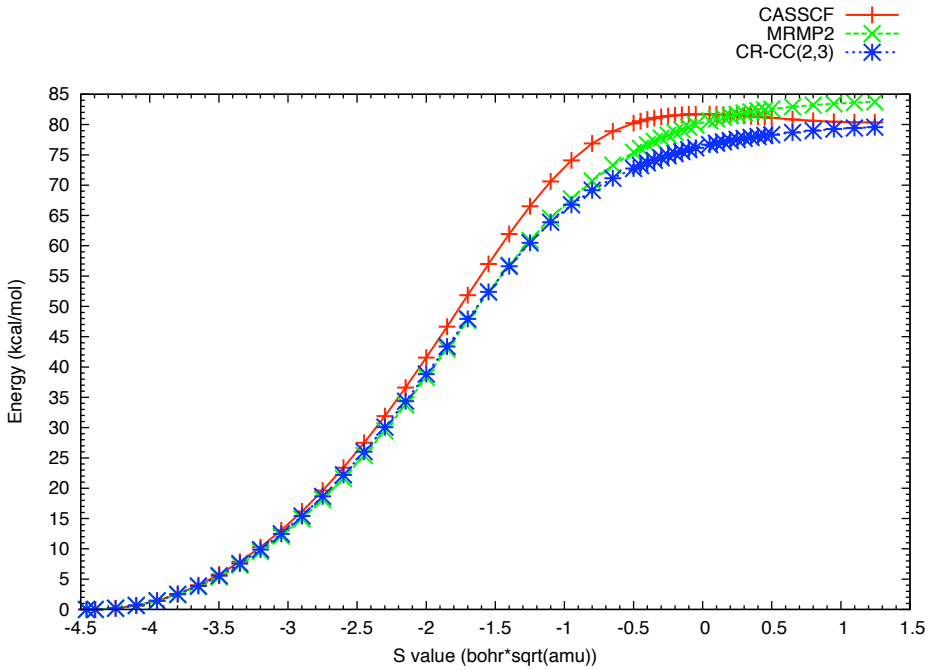


Figure 3. Path 22 Step 2: CASSCF IRC and Single Point Energies (aug-cc-pVTZ basis no ZPE)

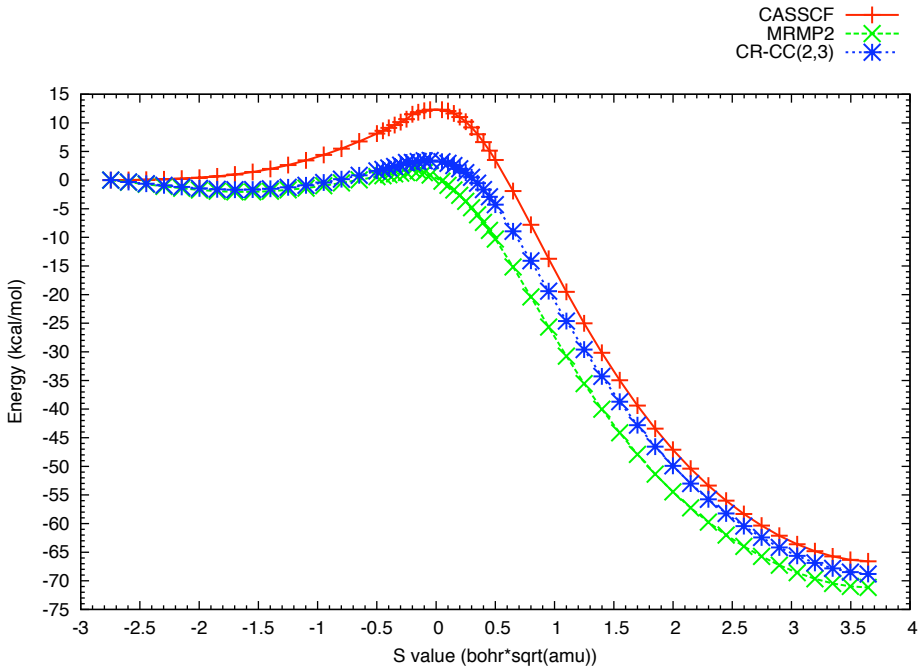
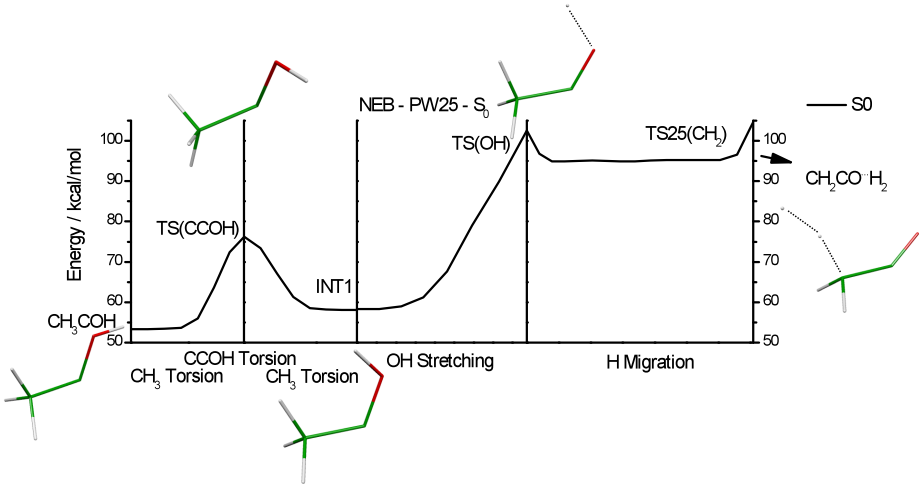


Figure 4. Non-traditional MEP for pathway 25 shown at the GVB-PP level with a cc-pVDZ basis. The plot displays several TSs between the desired minima.



References:

[1]West, A. C.; Kretchmer, J. S.; Sellner, B.; Park, K.; Hase, W. L.; Lischka, H.; Windus, T. L. *J. Phys. Chem. A* **2009**, *113*, 12663.

[2]West, A. C.; Lynch, J. D.; Sellner, B.; Lischka, H.; Hase, W. L.; Windus, T. L. *Theor Chem Acc* **2012**, *131*, 1123.

[3]Casavecchia, P.; Capozza, G.; Segoloni, E.; Leonori, F.; Balucani, N.; Volpi Gian, G. *J Phys Chem A* **2005**, *109*, 3527.

[4]Su, H.; Zhao, S.; Liu, K.; Xiang, T. *J Phys Chem A* **2007**, *111*, 9600.

[5]Lee, S.-H.; Huang, W.-J.; Chen, W.-K. *Chem. Phys. Lett.* **2007**, *446*, 276.

[6]Gardiner, W. C., Jr.; Editor *Combustion Chemistry*; Springer-Verlag, New York, 1984.

[7]Yamaguchi, K.; Yabushita, S.; Fueno, T.; Kato, S.; Morokuma, K. *Chem. Phys. Lett.* **1980**, *70*, 27.

[8]Dupuis, M.; Wendoloski, J. J.; Takada, T.; Lester, W. A., Jr. *J. Chem. Phys.* **1982**, *76*, 481.

[9]Fueno, T.; Takahara, Y.; Yamaguchi, K. *Chem. Phys. Lett.* **1990**, *167*, 291.

[10]Smith, B. J.; Nguyen Minh, T.; Bouma, W. J.; Radom, L. *J. Am. Chem. Soc.* **1991**, *113*, 6452.

[11]Jursic, B. S. *THEOCHEM* **1999**, *492*, 85.

[12]Nguyen, T. L.; Vereecken, L.; Hou, X. J.; Nguyen, M. T.; Peeters, J. *J. Phys. Chem. A* **2005**, *109*, 7489.

[13]Yang, X.; Maeda, S.; Ohno, K. *J. Phys. Chem. A* **2007**, *111*, 5099.

[14]Hu, W.; Lendvay, G.; Maiti, B.; Schatz, G. C. *J. Phys. Chem. A* **2008**, *112*, 2093.

[15]Joshi, A.; You, X.; Barckholtz Timothy, A.; Wang, H. *J Phys Chem A* **2005**, *109*, 8016.

[16]Shepler, B. C.; Braams, B. J.; Bowman, J. M. *J. Phys. Chem. A* **2008**, *112*, 9344.

[17]Heazlewood, B. R.; Jordan, M. J. T.; Kable, S. H.; Selby, T. M.; Osborn, D. L.; Shepler, B. C.; Braams, B.; Bowman, J. M. *Proc. Natl. Acad. Sci. U. S. A., Early Ed.* **2008**, 1.

[18]Shepler, B. C.; Braams, B. J.; Bowman, J. M. *J. Phys. Chem. A* **2007**, *111*, 8282.

[19]Gordon, M. S.; Schmidt, M. W. *Theory Appl. Comput. Chem.: First Forty Years* **2005**, 1167.

[20]Ruedenberg, K.; Sundberg, K. R. *Quantum Sci.* **1976**, 505.

[21]Cheung, L. M.; Sundberg, K. R.; Ruedenberg, K. *J. Am. Chem. Soc.* **1978**, *100*, 8024.

[22]Cheung, L. M.; Sundberg, K. R.; Ruedenberg, K. *Int. J. Quantum Chem.* **1979**, *16*, 1103.

[23]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M. *Chem. Phys.* **1982**, *71*, 51.

[24]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 65.

[25]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41.

[26]Feller, D. F.; Schmidt, M. W.; Ruedenberg, K. *J. Am. Chem. Soc.* **1982**, *104*, 960.

[27]Siegbahn, P.; Heiberg, A.; Roos, B.; Levy, B. *Phys. Scr.* **1980**, *21*, 323.

[28]Roos, B. *Adv. Chem. Phys.* **1987**, *69*, 399.

[29]Malmquist, P. A.; Olsen, J.; Taylor, P. R.; Roos, B. O.; Siegbahn, P. E. M.; Helgaker, T.; Wahlgren, U. *Lecture Notes in Quantum Chemistry, European Summer School in Quantum Chemistry. [In: Lect. Notes Chem., 1992; 58]*, 1992.

[30]Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

[31]Ivanic, J.; Ruedenberg, K. *Theor. Chem. Acc.* **2001**, *106*, 339.

[32]Lengsfield, B. H., III. *J. Chem. Phys.* **1980**, *73*, 382.

[33]Fletcher, G. D. *Mol. Phys.* **2007**, *105*, 2971.

[34]Yarkony, D. R. *Chem. Phys. Lett.* **1981**, *77*, 634.

[35]Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1980**, *72*, 880.

[36]Boys, S. F. *Quantum Theory At., Mol., Solid State. 1966* **1966**, 253.

[37]Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154.

[38]Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.

[39]Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett* **1993**, *214*, 481.

[40]Lischka, H.; Shepard, R.; Brown, F. B.; Shavitt, I. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1981**, *15*, 91.

[41]Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Muller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664.

[42]Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Mueller, T.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Boehm, H. J.; Chang, A.; Comeau, D. C.; Gdanitz, R.; Dachsel, H.; Ehrhardt, C.; Ernzerhof, M.; Hoechtl, P.; Irle, S.; Kedziora, G.; Kovar, T.; Parasuk, V.; Pepper, M. J. M.; Scharf, P.; Schiffer, H.; Schindler, M.; Schueler, M.; Seth, M.; Stahlberg, E. A.; Zhao, J.-G.; Yabushita, S.; Zhang, Z.; Barbatti, M.; Matsika, S.; Schuurmann, M.; Yarkony, D. R.; Brozell, S. R.; Beck, E. V.; Blaudeau, J.-P. **2006**, www.univie.ac.at/columbus.

[43]Piecuch, P.; Kucharski, S. A.; Kowalski, K.; Musial, M. *Comput. Phys. Commun.* **2002**, *149*, 71.

[44]Piecuch, P.; Wloch, M. *J. Chem. Phys.* **2005**, *123*, 224105/1.

[45]Hay, P. J.; Hunt, W. J.; Goddard, W. A., III. *J. Amer. Chem. Soc.* **1972**, *94*, 8293.

[46]Hunt, W. J.; Hay, P. J.; Goddard, W. A., III. *J. Chem. Phys.* **1972**, *57*, 738.

[47]Shepard, R.; Kedziora, G. S.; Lischka, H.; Shavitt, I.; Mueller, T.; Szalay, P. G.; Kallay, M.; Seth, M. *Chem. Phys.* **2008**, *349*, 37.

[48]Bunge, A. V. *J. Chem. Phys.* **1970**, *53*, 20.

[49]Pulay, P.; Fogarasi, G.; Pongor, G.; Boggs, J. E.; Vargha, A. *J. Am. Chem. Soc.* **1983**, *105*, 7037.

[50]Shepard, R. *Adv. Ser. Phys. Chem.* **1995**, *2*, 345.

[51]Shepard, R.; Lischka, H.; Szalay, P. G.; Kovar, T.; Ernzerhof, M. *J. Chem. Phys.* **1992**, *96*, 2085.

[52]Lischka, H.; Dallos, M.; Shepard, R. *Mol. Phys* **2002**, *100*, 1647.

[53]Mills, G.; Jonsson, H. *Phys. Rev. Lett.* **1994**, *72*, 1124.

[54]Mills, G.; Jonsson, H.; Schenter, G. K. *Surf. Sci.* **1995**, *324*, 305.

[55]Henkelman, G.; Jonsson, H. *J. Chem. Phys.* **2000**, *113*, 9978.

[56]Gonzalez-Garcia, N.; Pu, J.; Gonzalez-Lafont, A.; Lluch, J. M.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 895.

[57]Anglada, J. M.; Bofill, J. M. *J. Comput. Chem.* **1997**, *18*, 992.

[58]Fukui, K. *J. Phys. Chem.* **1970**, *74*, 4161.

[59]Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363.

[60]Sellner, B. *Dissertation, University of Vienna* **2011**.

[61]Chao, J.; Hall, K. R.; Marsh, K. N.; Wilhoit, R. C. *J. Phys. Chem. Ref. Data* **1986**, *15*, 1369.

[62]Holmes, J. L.; Lossing, F. P. *J. Am. Chem. Soc.* **1982**, *104*, 2648.

[63]Tsang, W. *Struct. Energ. React. Chem. Ser.* **1996**, *4*, 22.

[64]Niiranen, J. T.; Gutman, D.; Krasnoperov, L. N. *J. Phys. Chem.* **1992**, *96*, 5881.

[65]JANAF Thermochemical Tables; Natl. Stand. Ref. Data Ser.; U.S. National Bureau of Standards: Washington, D. **1985**, *37*.

[66]Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables, 4th ed.; J. Phys. Chem. Ref. Data, Mongr. 9; American Chemical Society: Washington, DC, 1998; p 1.*, *27*.

[67]Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[68]Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

[69]Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

[70]Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

[71]Chase, M. W., Jr.; Davies, C.A.; Downey, J.R., Jr.; Frurip, D.J.; McDonald, R.A.; Syverud, A.N., JANAF Thermochemical Tables (Third Edition), J. Phys. Chem. Ref. Data,Suppl. 1, 1985, 14, 1.

[72]Gurvich, L. V. V., I. V.; Alcock, C. B., Thermodynamic Properties of Individual Substances, Fouth Edition, Hemisphere Pub. Co., New York, 1989.

[73]Ruscic B, B. J., Burcat A,Csaszar AG,Demaison J, Janoschek R, Martin JML, Morton ML, Rossi MJ, Stanton JF, Szalay PG, Westmoreland PR, Zabel F, Berces T "IUPAC Critical Evaluation of Thermochemical Properties of Selected Radicals. Part I" J. Phys. Chem. Ref. Data, Vol. 34, No. 2, 2005, 573.

CHAPTER 5. CAN ORMAS BE USED FOR NONADIABATIC COUPLING CALCULATIONS? $SiCH_4$ AND BUTADIENE CONTOURS.

[With kind permission from Springer Science+Business Media: *Theor. Chem. Acc.,* Can ORMAS Be Used For Nonadiabatic Coupling Calculations? $SiCH_4$ and Butadiene Contours., 131, 2012, 1251, West, A. C.; Windus, T. L.;
1-20, Copyright Springer-Verlag 2012]

Aaron C. West and Theresa L. Windus

Abstract:

When appropriately used, the MCSCF approximation is useful in discerning correct electronic structure results. However, with the increasing size of chemical systems of interest, MCSCF rapidly becomes unfeasible due to the requirement of larger active spaces, which lead to computationally unmanageable numbers of configurations. This situation is especially true for CASSCF. In particular, reducing this computational expense by using restricted active spaces in solving for gradients and NACs during dynamics runs would save computer time. However, the validity of such restricted spaces is not well-known even for recovering the majority of the nondynamical correlation and will inevitably vary between chemical systems across a range of nuclear geometries. As such, we use the recently implemented CP-ORMAS equations[1] to verify the accuracy of this approximation for gradients and NAC vectors around two specific C.I. geometries for the silaethylene and butadiene systems. Overall, no excitations between appropriate subspaces show qualitatively

reasonable results while single excitations significantly improve ORMAS results relative to the CASSCF level in these particular systems. However, single excitation schemes do not always lead to the correct orbital subspaces, and as a result, seemingly smooth PES do not always result in smooth analytical gradients and NACMEs. In addition, while some of the single excitation ORMAS and CASSCF schemes have improper orbitals rotate into the active space, the schemes without excitations (even with more subspaces) do not exhibit this behavior.

Introduction:

Creating computationally inexpensive and reliable wavefunction approximations remains a difficult task. To date, multiconfigurational self-consistent field (MCSCF) with a sufficiently flexible active space (as defined in reference [2]) proves to be a reliable method for generally mapping a potential energy surface (PES). In other words, if an active space choice fails to produce the qualitatively correct PES features, it is insufficiently flexible. However, sometimes the active space may have to be quite large to be sufficiently flexible enough to simultaneously map several reaction pathways of interest. The most commonly employed MCSCF theory level remains the full optimized reaction space (FORS),[3-9] or complete active space self-consistent field (CASSCF).[10-12] However, as chemical systems of interest increase in number of required active orbitals, CASSCF configuration interaction (CI) expansion sizes become unwieldy, and CASSCF calculations quickly become intractable since a linear increase in the number of active orbitals results in a factorial increase in the number of configurations. Due to this expense, CASSCF computations tend to contain small active spaces and suffer energetic inaccuracy from lack of correlation. Nonetheless, recovering the majority of nondynamical correlation with fewer configurations

still remains a relevant topic. Occupation restricted multiple active space (ORMAS)[13,14] provides one alternative MCSCF method to decrease the CI expansion size and still possibly maintain accuracy in mapping a PES. The ORMAS method is similar to several other CI expansion size reduction methods: restricted active space,[15] restricted configuration interaction,[16,17] quasi-complete active space,[18] general active space,[19] and macroconfigurational[20] methods. In the ORMAS approximation, excitation restrictions between different invariant orbital subspaces reduce the overall CI expansion size in an *a priori* fashion. Constructing ORMAS wavefunctions consists of partitioning a total active space into several subspaces and defining maximum and minimum electron occupation values for each subspace that in turn represent excitations of electrons between the subspaces. The CI and orbital coefficients are then self-consistently optimized. If the excitations between subspaces are limited, there is the potential for large computational savings (see Methods section). However, selecting both the orbital subspaces and the allowed excitations between them requires both chemical intuition and numerical testing.

While ORMAS has been used in PES mapping, this approximation can also now be applied to ground as well as excited state analytical gradient techniques[21-24] and nonadiabatic coupling matrix element (NACME)[25-27] formation. For analytical gradient and NACME equations, Lengsfield et al.[26] give the general MCSCF implementation, and Lischka et al.[27] convey the equations for the multireference configuration interaction (MR-CI) level that allow for the general selection of MR-CI spaces. In addition for multireference perturbation (MR-PT), analytical gradients are established for other post-MCSCF methods such as second-order complete active space perturbation theory (CASPT2)[28] and canonical quasi-degenerate perturbation theory (QDPT).[29] Also, analytical NACMEs are derived and

implemented for the generalized Van Vleck perturbation theory[30,31]. Approximations that reduce a CI expansion length also reduce analytical gradient and nonadiabatic computation times, which in turn reduce times for nonadiabatic dynamics runs and conical intersection (C.I.) [32-39] searches.

Obtaining accurate results with approximations such as ORMAS is not always well understood, and to the best of our knowledge, no studies on the more demanding criteria of ORMAS gradient and NACME accuracy have been conducted. *In fact, many studies that rely on incomplete active space results access accuracy based on either energy continuity or energy conservation criterion alone, and such energy assessments alone can actually fail to produce smooth analytical gradients and NACMEs.* In order to give some justification for the use of ORMAS gradients and NACMEs, this study examines the C.I.s – where NACMEs are usually large – and nearby contours for silaethylene and butadiene with limited ORMAS and complete active space (CAS) sizes. This limitation is employed because the expense of CASSCF gradients and NACMEs prevents comparisons with larger active spaces.

Several recent investigations on ethylene[40,41] and silaethylene[42,43] examined the excited states in great detail with calculations at various levels of theory, some of which include MR-CI. Silaethylene investigations are interesting because of the very few excited state studies on the $SiCH_4$ polarized π bond. In marked contrast to ethylene, the SiC torsion from the planar to the orthogonal H_2SiCH_2 structure leads to a C.I., and the $\pi\pi^*$ excited state is the lowest excited singlet state and is well separated from the other excited singlet states. Furthermore, in contrast to some static structure studies, recent dynamics studies suggest pyramidalizations are involved in C.I.s as well. Butadiene investigations[35,36,44-47] are more extensive with controversy over the mechanisms. Olivucci et al.[44] disputed several reaction

mechanisms on the s-cis/s-trans isomerization pathway and provided one of the earliest papers on butadiene C.I.s, which include the s-cisoid, central, and s-transoid C.I. Later studies[46] on butadiene focused more on solving the disputed ordering of singlet states with extremely accurate energetics.

This study does not seek to answer any of these reaction mechanism issues because it is limited to MCSCF wavefunctions. Instead, this study focuses on the nondynamical correlation comparison of ORMAS and CASSCF. While agreement between these levels of theory is beneficial, some disagreement does not prevent insights into how variance of nondynamical MCSCF correlation affects gradients and NACMEs. In addition, as long as the ORMAS results can reasonably describe the nondynamical correlation effects relative to the CASSCF results, the ORMAS wavefunctions may still be very useful as reference wavefunctions in post-MCSCF methods. Knowledge of any limitations at the MCSCF level will prove useful in both future applications and methods development.

Methods:

All calculations were performed with the GAMESS software.[48] State averaged[49] (SA)-MCSCF energy calculations were carried out with equal SA weights at the CASSCF[11,12] and the ORMAS[13,14] levels of theory for the lowest-lying singlet states of interest. For the ORMAS wavefunctions in particular, the determinant-based method and the Jacobi[50] and full Newton-Raphson convergers with augmented Hessian technique were utilized.[51-53] The 6-31G(d)[54-57] basis set was used for these comparisons. In addition, for one of the active spaces for the $SiCH_4$ system the 6-311++G(d,p)[58-60] basis set was also used in order to demonstrate that these results show little to no basis set dependence. However, basis set dependence in the analysis could be an issue for other chemical systems and will be

explored in future work. Following an energy run, state-specific gradients (simply referred to as gradients) and NACMEs were formed as directed by Lengsfield[26] and as implemented in Dudley et al.[61] (The coupled perturbed (CP)-ORMAS implementation is currently unpublished.[1]) In order to locate PES regions with large nonadiabatic couplings (NACs) for the comparisons, SA-CASSCF minimum energy C.I. geometry searches were first performed by way of gradient projection with NACME.[39] (For brevity the minimum energy C.I. is referred to as C.I. below.) Once a C.I. geometry was located, rigid changes in two internal degrees of freedom were performed in order to obtain some variation in the energies, gradients, and NACs.

For the $SiCH_4$ system in this study, the two state SA-2-(4,4)CAS optimized C.I. geometry was identified as the CH_2-pyramidalized twisted orthogonal C.I.[42] As indicated in Pitonak and Lischka,[42] the energy difference gradient vector direction largely coincides with the SiC stretch direction, and the NAC vector direction largely coincides with the direction of HSiCH dihedral rotation. As such, various ORMAS schemes and a CAS scheme of SA-2 single point energies followed by gradients and NACMEs were performed on a grid where the SiC distance was rigidly increased from 1.5 – 2.2 Å by 0.05 Å increments and the HSiCH dihedral angle was rigidly rotated from -75° to -55° by 1° increments. The initial silaethylene contour calculations were done with a SA-2-(4,4)CAS, which contained SiC σ, σ*, π, π* in the active space. For the ORMAS runs, three schemes were used. Scheme 1 was chosen as SA-2-2(2,2)ORMAS with no excitations between the two separate SiC σ, σ* and SiC π, π* spaces. Scheme 2 was identical to scheme 1 but with single excitations between spaces. Further ORMAS testing verified that including double excitations gives the equivalent CASSCF results. The third scheme (i.e. the 'occupied' scheme) was SA-2-2(2,2) ORMAS

with single excitations between SiC σ, π and SiC σ*, π* spaces. In order to report consistent active space results with one particular set of orbitals (i.e. one analytical surface) for the third scheme, any scheme 3 contours are restricted to a SiC distance of 1.5 – 1.8 Å. Outside of this domain, the ORMAS energies lie on a different analytical surface.

Because different analytical surfaces can intersect[2] and in some cases still produce seemingly smooth PESs, ensuring all geometries have proper active space orbitals can prove difficult in any grid study. In order to demonstrate correctness of the $SiCH_4$ results, more C.I.s and contours were formed but with a larger (8,8) active space and the 6-31G(d) and 6-311++G(d,p) basis sets. For the calculations with the 6-311++G(d,p) basis set, the same contour grid was formed as before, but for the calculations with the 6-31G(d) basis set, a smaller, finer grid was used from 1.6 – 1.9 Å and from -75^o to -55^o. Furthermore, for this finer grid, the contour angle increments were kept as 1^o while 0.01 Å increments were used for the SiC distance. Then, the same smaller grid can also be formed with 0.05 Å increments for comparison. Formed from these two different distance increments, the two grids indicate 0.05 Å increments are sufficient based on comparisons of the results. Of course, peaks for the finer grid are larger since the NACMEs diverge near the C.I., but the overall conclusions remain the same. The initial silaethylene contours were created with a SA-2-(8,8)CAS, which contained SiC σ, σ*, π, π* and two SiH σ, σ* orbitals in the active space. The SiH orbitals were added to prevent orbital problems in the SA-2(4,4)CAS. For the ORMAS runs, two schemes were used. Scheme 1b was chosen as a SA-2-4(2,2)ORMAS with no excitations between four separate spaces: SiC σ, σ*, SiC π, π*, SiH σ, σ*, and SiH σ, σ*. Scheme 2b was identical to scheme 1b but with single excitations between the four spaces. (Scheme 2b calculations were performed only at the 6-311++G(d,p) level.)

For the butadiene system, the SA-2-(10,10)CAS optimized C.I. geometry was identified as the s-transoid C.I.[44] As discerned in Olivucci et al.,[44] central CC rotation and "asynchronous disrotatory rotation of the two terminal methylenes" partially reveal the C.I. curvature. So, SA-2-(10,10)CAS single point energies, gradients, and NACMEs were performed on a grid by rigid rotations on the CCCC dihedral from 114^{o} to 122^{o} by 1^{o} increments and on the one CCCH dihedral from -105^{o} to -95^{o} by 1^{o} increments. For these initial CAS single points, the active space includes all of the CC π, π^*, σ, and σ^* orbitals. Then, two ORMAS schemes were performed over the same orbital range: (4,4) + 3(2,2)ORMAS with one CC π, π^*, π, π^* space and three CC σ, σ^* spaces with no excitations and with single excitations between spaces. The SA-2-(4,4)+3(2,2)ORMAS space with no and single excitations will simply be referred to as ORMAS(butadiene1) and ORMAS(butadiene2) below, respectively.

Overall, given the use of a determinant-based implementation, the numbers of determinants in the calculations are useful for understanding the computational expense for each MCSCF level of theory. For the $SiCH_4$ system based off of the (4,4) active space size, the total numbers of determinants for each scheme are 36, 18, 34, and 34 for (4,4)CAS, scheme 1, scheme 2, and scheme 3, respectively. For the $SiCH_4$ system based off of the (8,8) active space size, the total numbers of determinants for each scheme are 4900, 454, and 3526 for (8,8)CAS, scheme 1b, and scheme 2b, respectively. For the butadiene system, the total numbers of determinants for each scheme are 63504, 4824, and 41256 for (10,10)CAS, ORMAS(butadiene1), and ORMAS(butadiene2), respectively.

Results and Discussion:

All C.I. geometries are reported at the CASSCF level of theory and are used as a reference for the subsequent contours. Contour plots about the C.I. geometries are presented in Figures 1 - 6. Additional contours can be found in the Supporting Information, which includes additional plots for energies, gradients, and NACMEs. Phrases like "C.I. contour" and variations on it simply indicate the contour data comes from the grids discussed in the Methods section. The "total gradient magnitude" and "total NACME magnitude" values equal the square root of the sum of the squares of the gradient and NAC vector components, respectively. Any phrase with difference(s) or "magnitude difference(s)" simply means the subtractive difference between two different theory level results of interest. In order to compare CASSCF and ORMAS gradient or NACME vectors, overlap (i.e. multi-dimensional dot products) plots display the directional differences between normalized vectors at these two different levels of theory with a maximum absolute value overlap value of 1. Right next to the CASSCF C.I. geometry, single points at the ORMAS theory level can shift the curvature, and in the case of butadiene, these shifts result in 2-3 negative gradient overlaps. Unless the different levels of theory produce exactly identical PESs (i.e. no shift in the location of the C.I. occurs), this behavior is to be expected in the geometric vicinity of the C.I., which has drastic changes to the state gradients. So, given the slightest C.I. shift from single point runs at different levels of theory, the corresponding gradient or NAC vectors can show different directions and result in poor overlap. Decreasing the contour increment sizes will reveal more poor overlaps; however, even overlaps only slightly farther from the C.I. geometry typically show excellent overlaps. So, in the case of butadiene, the absolute values

of the gradient overlaps are taken at the previously noted 2-3 geometries for more clarity in the plots.

$SiCH_4$:

For the $SiCH_4$ case, the SA-2-(4,4)CAS/6-31G(d) CH_2-pyramidalized twisted orthogonal C.I. was located. The only appreciable difference found between the C.I. geometries in this study and the Pitonak and Lischka C.I. geometry[42], which is at the SA-3-CAS(2,2)/MR-CISD/aug-cc-pVDZ level of theory, is in the CH_2 pyramidalization found in this study. Our calculations show that this pyramidalization could be associated with the larger active space. The C.I. geometry approximately occurs at a SiC distance of 1.73Å and with one HSiCH dihedral at 74^o, and as shown in the Pitonak and Lischka study, very slight CH_2-pyramidalization leaves the lowest-lying states almost degenerate. Figure 1 gives the $SiCH_4$ singlet state energy contours at the SA-2-(4,4)CAS/6-31G(d) level where all energies are plotted relative to the lowest energy in the lowest-lying singlet energy contour. Figure 2 gives the excitation energy differences between the two states at the same level of theory, and Figure 3 shows the excitation energy difference contour at the ORMAS(scheme 1)/6-31G(d) level. The Figure 4 contour displays the difference between the (4,4)CAS and ORMAS(scheme 1) excitation energies and thus explicitly demonstrates the energetic accuracy of ORMAS(scheme 1) very near the given CASSCF C.I. geometry. However the excitation energies significantly deviate farther away from the C.I.

Figures 5 and 6 show that the total NACME magnitude locally peaks around the same nuclear geometries for both the CASSCF and the ORMAS(scheme 1) approximations. Figure 6 shows a spike of 56 $Bohr^{-1}$ in the differences for the total NACME magnitudes. However, it is important to note that at the C.I. geometry itself, the CAS total NACME

magnitude actually approaches ~1x10^{5} Bohr^{-1} (about four orders of magnitude larger than those magnitudes shown in the Figures) with the C.I. point only ~0.02 Å from the closest surrounding grid points. The plots do not contain the C.I. itself since then the NACME magnitude contours would be flat with one point as a peak. This observation emphasizes that the NACME magnitude spikes in an extremely local fashion. Since the grid points do not lie on the exact peak at either level of theory, the NACME magnitudes comparisons are actually quite good. However, the (8,8) active space results below will show subtle differences, which can possibly be attributed to improper orbitals in some geometries near the C.I. in the (4,4) active space. Figure 7 shows the NACME overlaps between the (4,4)CAS and ORMAS(scheme 1). Very close to the CASSCF C.I. geometry, the CAS and ORMAS NAC overlap is very poor for two points. However, NAC directionality between the two levels of theory approaches a total overlap of 1 at most geometries.

ORMAS(scheme 1) shows some differences in the gradients relative to the (4,4)CAS level of theory. The largest gradient magnitude difference for state 1 (2) gradient is ~|0.1| (|0.06|) hartrees/Bohr. However, the trend for the gradient normalized vector overlaps is just the opposite. For state 1 the ORMAS(scheme 1) gradient overlaps with the CAS gradient by > 90% at all points. However, for state 2 the ORMAS(scheme 1) gradient shows little overlap at a few geometries very near to the C.I. (Figure 8). At least part of this poor overlap reflects differences in the location of the C.I. between the two levels of theory. Farther away from the C.I. geometry, the overlaps approach a value of 1, which indicates the two levels of theory qualitatively produce very similar gradients.

For ORMAS(scheme 2) results relative to the SA-2-(4,4)CAS results, the energy differences are less than |0.05| kcal/mol, and the total NACME magnitude differences are

well within |1| Bohr^{-1} at all points on the contours (see Supporting Information for the plots for schemes 2 – 3). The total gradient magnitude difference contours reveal some differences on the order of ~1x10^{-2} hartrees/Bohr. For the state 1 gradient overlaps, CASSCF and ORMAS gradient vectors always overlap by >95%. State 2 gradient overlap shows the worst overlap of ~70% at a geometry very near the C.I. (Figure 9). For overlaps, ORMAS(scheme 2) values are in much better agreement with CASSCF than ORMAS(scheme 1). All the plotted NAC vectors between states 1 and 2 also have >92% overlap. ORMAS(scheme 3) and CASSCF energy and NACME magnitude comparisons reveal the same essential pattern as in scheme 2 with very slight differences (i.e. |0.03| kcal/mol, |3| Bohr^{-1}, respectively). Again, state 1 gradient overlaps are >96%, and NAC overlaps also show >94% overlap. State 2 gradient overlap (Figure 10) shows poor overlaps of ~40% at a few geometries.

For additional comparison, the CH_2-pyramidalized twisted orthogonal C.I. was located at the SA-2-(8,8)CAS/6-31G(d) and SA-2-(8,8)CAS/6-311++G(d,p) levels, and the corresponding contours were constructed. For the C.I. geometry, no appreciable differences were found between the SA-2-(4,4)CAS/6-31G(d) and SA-2-(8,8)CAS/6-311++G(d,p) levels of theory. For the SA-2-(8,8)CAS and ORMAS(scheme 1b) levels with the 6-311++G(d,p) basis, the total NACME magnitudes have a maximum difference of |22| Bohr^{-1} (Figure 11) in contrast to the |56| Bohr^{-1} peak difference in the (4,4) results shown in Figure 6. Upon further investigation, the larger NACME magnitude difference for the (4,4) results occurs because of improper active spaces at geometries near the C.I. for the (4,4)CAS. For these geometries, the variational principal prevents proper active spaces. However, this comparison now strengthens the results by showing the NACME magnitude maxima agree

and by revealing any possible problem areas with the (4,4) active space results by comparison.

For the SA-2-(8,8)CAS/6-311++G(d,p) and ORMAS(scheme 1b)/6-311++G(d,p) levels (collectively referred to as the (8,8) level from now on for difference contours), Figure 11 shows the two, almost symmetric peaks for the total NACME magnitude difference contour. To emphasize that the NACME peaks do not occur at the same geometry between the two levels of theory, the SA-2-(8,8)CAS (ORMAS(scheme 1b)) NACME magnitude peak resides at a SiC distance of 1.75 (1.65) Å and a HSiCH dihedral angle of -67 (-67)^{o}. In addition, the (8,8) NACME magnitude peak heights for CAS (27 Bohr^{-1}) and ORMAS (22 Bohr^{-1}) are very similar in contrast to the corresponding (22 and 78 Bohr^{-1}) (4,4) peak heights. The (8,8) total gradient magnitude difference contours for state 1 (Figure 12) and state 2 (Figure 13) also nicely mirror each other about the C.I. inflection point. The (8,8) state 1 gradient overlaps remain above 90% while the state 2 gradient overlaps show a minimum overlap of ~30%, which is a much lower overlap relative to the corresponding (4,4) results. As well, the (8,8) NAC overlap plot still shows a low overlap near the C.I. and more points with poor overlaps (Figure 14). However, as shown by the (8,8) active space/6-31G(d) levels in the Supporting Information, these additional poor overlaps do not seem to result from increasing the basis set size but from increasing the active space size given the same contour increments.

For the SA-2-(8,8)CAS/6-311++G(d,p) and ORMAS(scheme 2b)/ 6-311++G(d,p) levels, the single excitation results tend to match the CASSCF results. The following results are for the 1.5 to 1.9 Å SiC distance range, which display good analytical gradient and NACME results despite incorrect active orbital subspaces at many geometries. Indeed, even

though the entire active space contains the proper orbitals, the active orbital subspaces are incorrect and will not precisely generate the previously expected configurations/determinants. Nonetheless, the generated configurations essentially reproduce the CASSCF results at many geometries. All excitation energy differences are < |0.6| kcal/mol, and all vector overlaps are >90%. All gradient magnitude differences are < |0.01| hartrees/Bohr, and all NACME magnitude differences are < |8| $Bohr^{-1}$. Outside of this range far away from the C.I., a few geometries show discontinuities in the energy difference contour, and one geometry shows terrible overlaps of ~80% due to active orbital subspace problems (Supporting Information). Despite these issues, ORMAS(scheme 2b) still results in contours with smooth energetics over the majority of the 1.5 to 2.2 Å SiC distance range and smooth gradient and NACME magnitudes and overlaps over the 1.5 to 1.9 Å SiC distance range.

Overall for this case, no excitations between the ORMAS subspaces (scheme 1 and 1b) yield qualitatively reasonable NACMEs while at least single excitations are required to give reasonable energies and gradients relative to the CASSCF. In addition, smooth energy contour plots do not indicate proper orbital subspaces over the range of geometries, and even smooth gradient and NACME vectors are not definitive proof of the orbital subspace integrity. However, gradient and NACME discontinuities can occur even with smooth energy contours. In these cases, these analytical derivative discontinuities indicate an incorrect active space.

Butadiene:

In addition to the spaces described in the Methods section, a SA-2-(4,4)CAS and SA-2-2(2,2) ORMAS space with no excitations between two separate CC π, π^* spaces were

attempted. However, for the given geometry, excitations between the π spaces appear to be very important, and the ORMAS PES is not comparable to the (4,4)CAS PES around the (4,4)CAS C.I. This notion makes sense given that at the twisted geometries, increasing the allowed excitations results in additional, energetically important configurations to enter the wavefunction from the 'united' CC π, π* spaces. As previously mentioned, the ORMAS space must give a good representation of the chemistry and the PES before it can be used for other purposes.

For the butadiene case, the SA-2-(10,10)CAS/6-31G(d) C.I. was located. Some appreciable differences are found between this and the Olivucci et al.[44] C.I., which is at the SA-(4,4)CAS/4-31G level of theory. This (Olivucci[44]) study reports the terminal HCCH dihedrals as 34^o (24^o) and 73^o (78^o). Figure 15 gives the differences between ORMAS(butadiene1) and (10,10)CAS excitation energies and shows a maximum magnitude of ~|2.0| kcal/mol.

Figure 16 demonstrates that the ORMAS(butadiene1) NACME approximation remains valid relative to the CASSCF NACME in the butadiene case in terms of total magnitudes and also shows how the NACs peak at only slightly different geometries. The ORMAS(butadiene1) total gradient magnitude difference contours give a maximum difference on the order of $\sim 1 \times 10^{-1}$ hartrees/Bohr close to the C.I. Figures 17-19 show slightly more geometries with poor overlaps relative to the $SiCH_4$ (4,4) results. However, most geometries still show >90% overlap. These results represent qualitatively good agreement between the two levels of theory.

For the ORMAS(butadiene2) case, single excitations further improve and affirm the validity of the ORMAS approximation in the construction of NACMEs and gradients. The

ORMAS(butadiene2) and (10,10)CASSCF excitation energies differ by a maximum of |0.5| kcal/mol. Also for these two levels of theory, the total NAC magnitude difference maximum is now |60| $Bohr^{-1}$, and both state 1 and state 2 total gradient magnitude difference contour maximums decrease to ~|$5x10^{-2}$| hartrees/Bohr. Furthermore, both gradients and the NACMEs now have >84% overlap at all geometries on the grid except for a single geometry. Figure 20 shows the ORMAS(butadiene2) and (10,10)CAS NACME vector overlaps with a lower overlap of ~75% only at a single geometry.

Conclusions:

Based on the results from the $SiCH_4$ and butadiene systems at specific C.I. geometries with carefully chosen CAS and ORMAS active spaces, ORMAS single excitation schemes can often produce CASSCF-like results but tend to have orbital subspace integrity issues, which have the potential to get worse as the subspace sizes increase. On the other hand, schemes with no excitations between subspaces appear to give no orbital subspace issues at the same geometries. As a possible consequence, choosing the smallest, most chemically reasonable orbital subspaces, performing ORMAS with no excitations between those subspaces, and then applying a chosen MR-CI excitation scheme might help to eliminate many orbital issues from calculations. However, for the larger subspaces with the procedure above, the nondynamical correlation lost should be determined and characterized.

Also based on the same specific systems, this study characterizes the qualitative accuracy of ORMAS NACME in terms of total magnitudes and additionally indicates that CAS and ORMAS NAC vectors strongly coincide at most geometries. For these particular systems, the total NACME magnitudes are extremely localized for both CAS and ORMAS, and the largest NACME magnitudes occur close to one another but not at the same geometry.

The NACME magnitudes are very sensitive to the proximity of the C.I., and ORMAS gives the appropriate behavior near the CASSCF C.I. Because the ORMAS level of theory can shift the maximum magnitudes, the contours also reveal the differences between CAS and ORMAS total gradient magnitudes (i.e. $\sim 10^{-2} - 10^{-1}$ hartrees/Bohr) about the C.I. geometry with modest energy differences (i.e. 1 – 2 kcal/mol). However, other systems might have several C.I.s in geometrical proximity and thus have more reactivity with 'consecutive', minimum energy C.I.s. Such an idea would complicate the overlaps results. Further investigation is required into C.I.s at the ORMAS level of theory (i.e. C.I.s versus avoided crossings with varying theory level) and how ORMAS NACME affects product distributions during dynamics.

Acknowledgments:

The authors are indebted to Michael W. Schmidt and Mark S. Gordon for help in using the capabilities of GAMESS and MCSCF. This material is based upon work supported by the National Science Foundation under Grant No. OISE-0730114 for the Partnerships in International Research and Education (PIRE). TLW acknowledges computing resources purchased through funds provided by Ames Laboratory and Iowa State University.

Electronic Supplemental Material:

The online version of this article (doi:10.1007/s00214-012-1251-6) contains supplementary material, which is available to authorized users.

Figure 1. $SiCH_4$ C.I. (4,4)CAS Energy (kcal/mol) Contours SA-2/6-31G(d) Level.

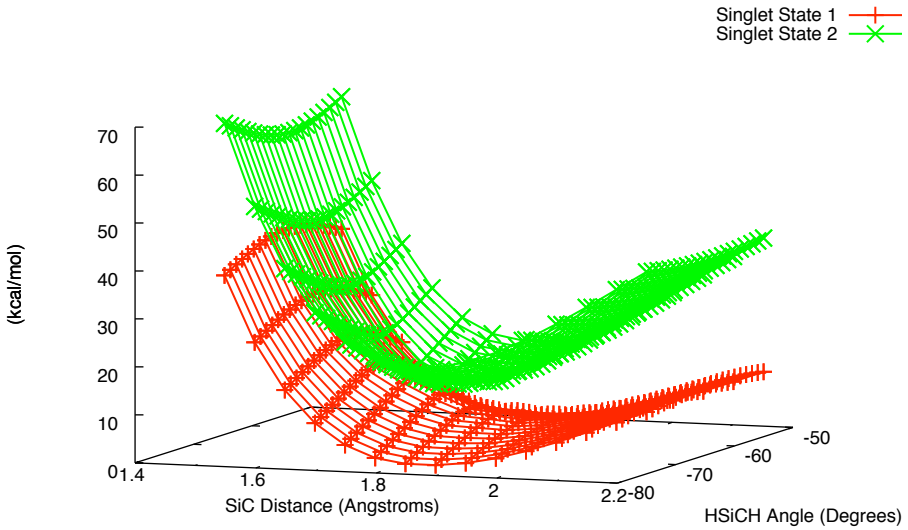


Figure 2. $SiCH_4$ C.I. (4,4)CAS Excitation Energy (kcal/mol) Contours SA-2/6-31G(d) Level.

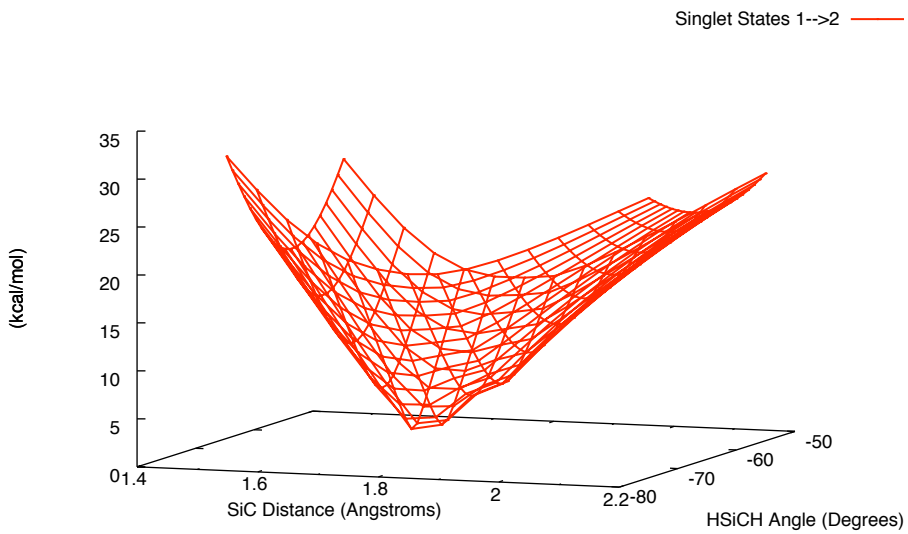


Figure 3. $SiCH_4$ C.I. ORMAS(scheme1) Excitation Energy (kcal/mol) Contours SA-2/6-31G(d) Level.

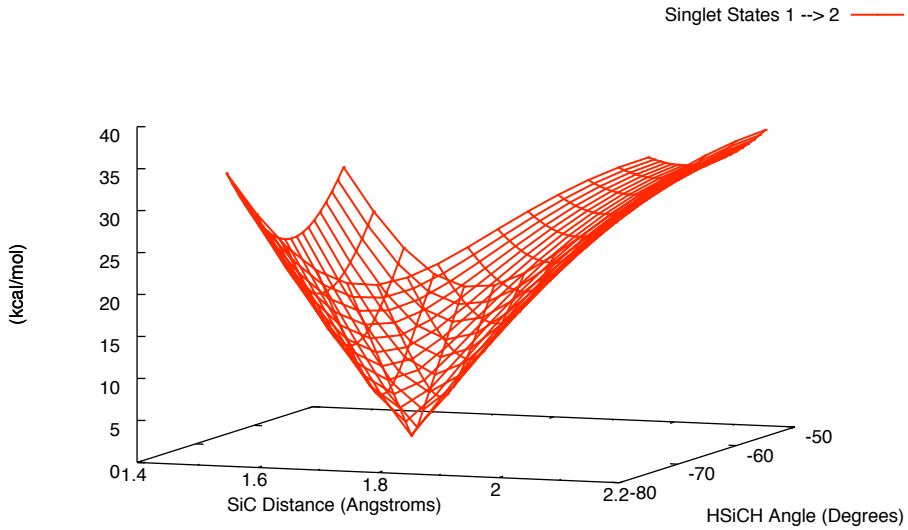


Figure 4. $SiCH_4$ C.I. ORMAS(scheme 1) – (4,4)CAS Excitation Energy (kcal/mol) Difference Contours SA-2/6-31G(d) Level.

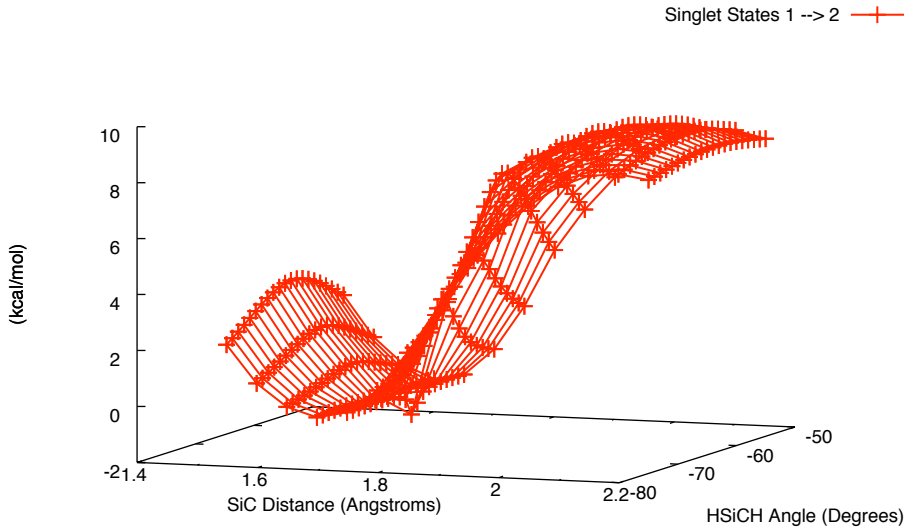


Figure 5. $SiCH_4$ C.I. (4,4)CAS Total Magnitude ($Bohr^{-1}$) NACME Contours SA-2/6-31G(d) Level.

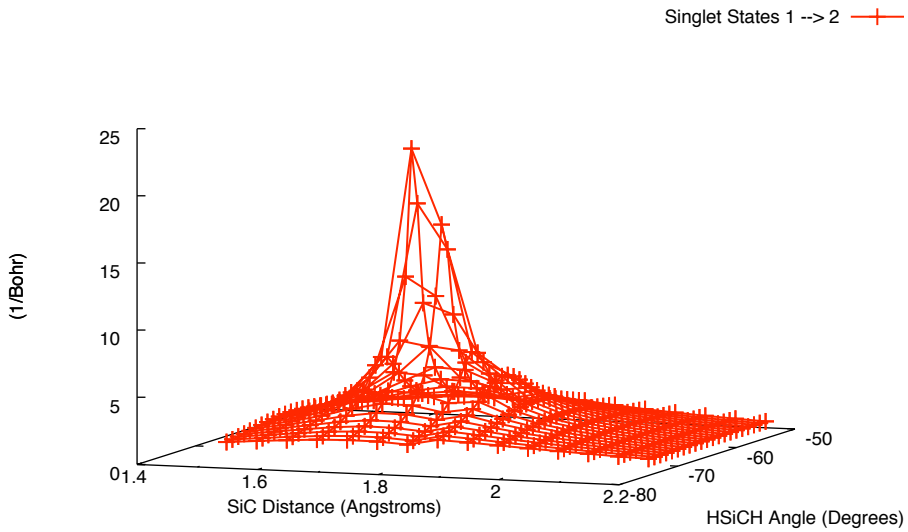


Figure 6. $SiCH_4$ C.I. ORMAS(scheme 1) – (4,4)CAS Total Magnitude ($Bohr^{-1}$) NACME Difference Contours SA-2/6-31G(d) Level.

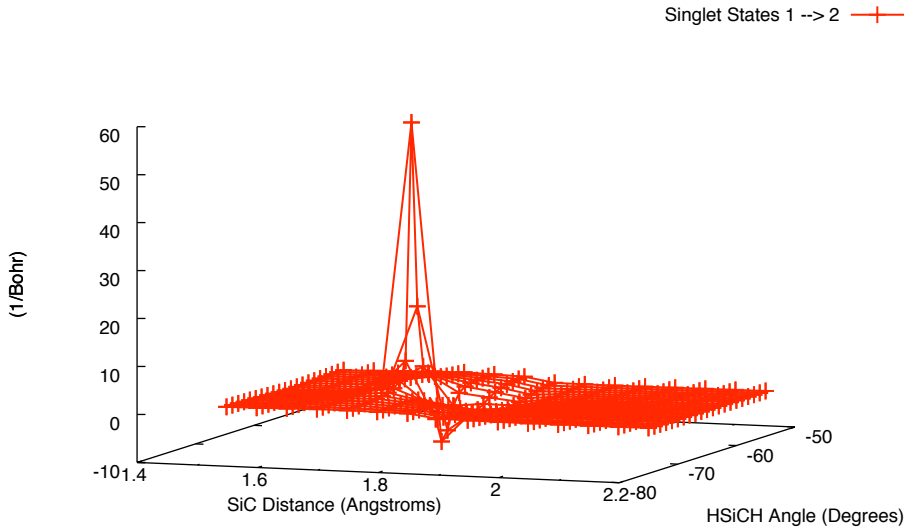


Figure 7. $SiCH_4$ C.I. ORMAS(scheme 1) and (4,4)CAS NACME Normalized Vector Overlap Contours SA-2/6-31G(d) Level.

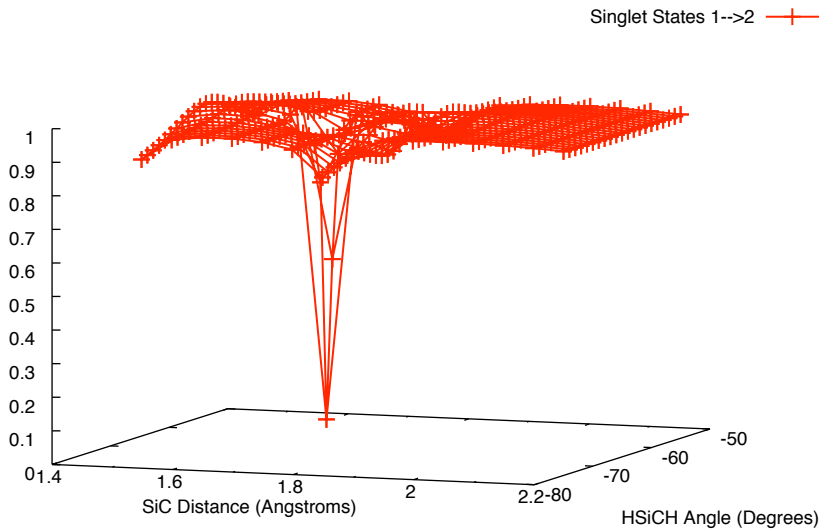


Figure 8. $SiCH_4$ C.I. ORMAS(scheme 1) and (4,4)CAS State 2 Gradient Normalized Vector Overlap Contours SA-2/6-31G(d) Level**.**

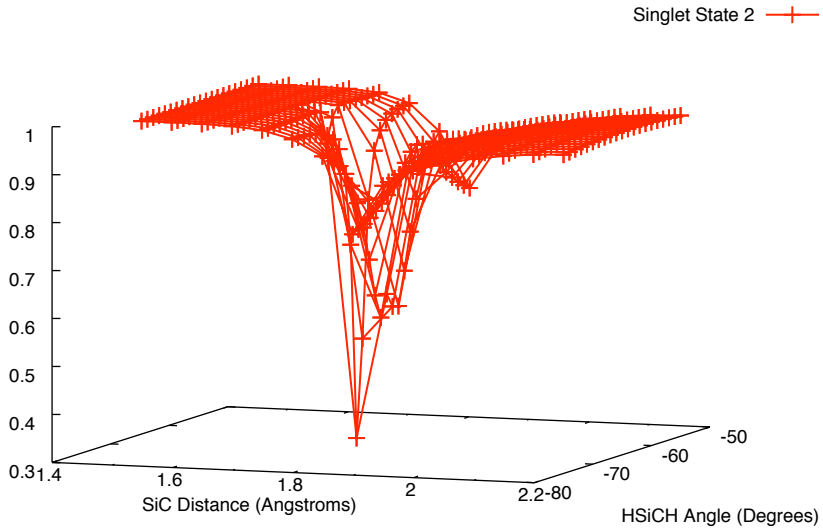


Figure 9. $SiCH_4$ C.I. ORMAS(scheme 2) and (4,4)CAS State 2 Gradient Normalized Vector Overlap Contours SA-2/6-31G(d) Level**.**

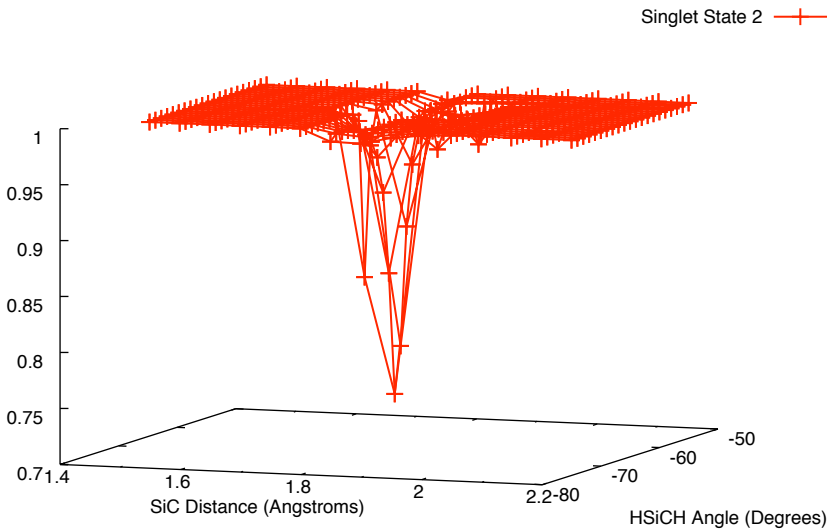


Figure 10. $SiCH_4$ C.I. ORMAS(scheme 3) and (4,4)CAS State 2 Gradient Normalized Vector Overlap Contours SA-2/6-31G(d) Level**.**

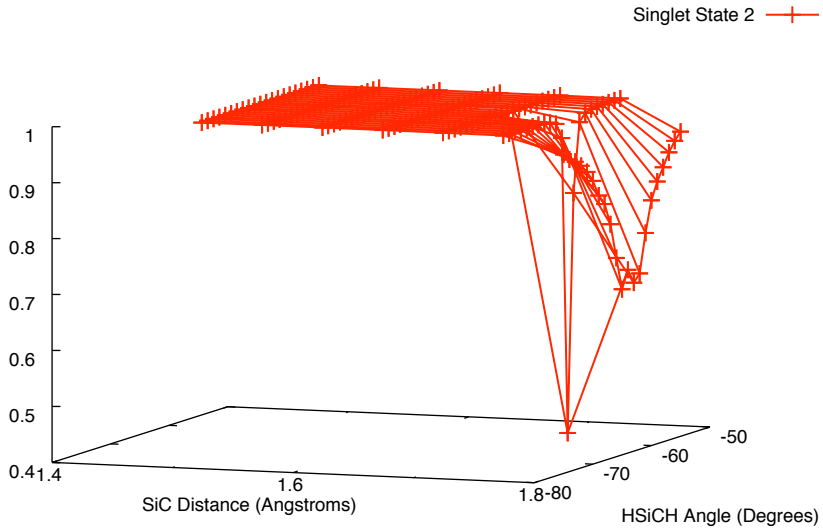


Figure 11. $SiCH_4$ C.I. (ORMAS(scheme 1b) – (8,8)CAS) Total NACME Magnitude ($Bohr^{-1}$) Difference Contours SA-2/6-311++G(d,p) Level.

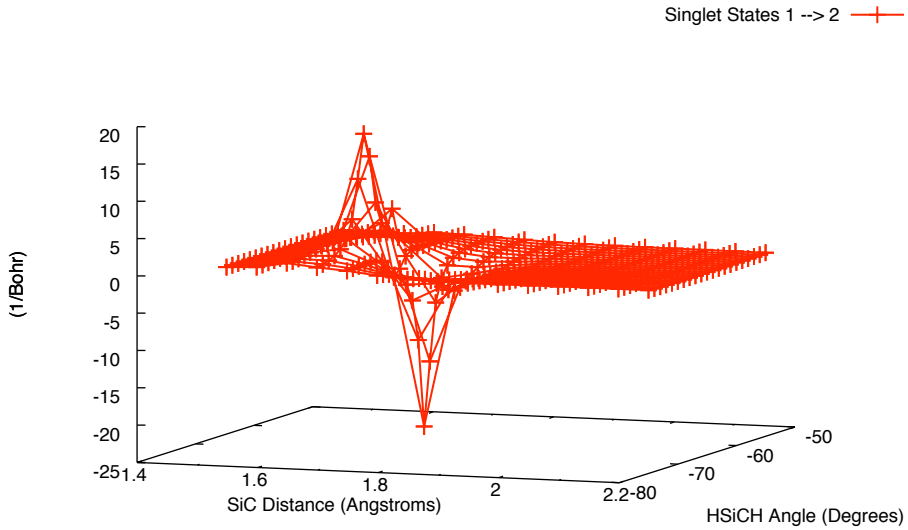


Figure 12. $SiCH_4$ C.I. (ORMAS(scheme 1b) – (8,8)CAS) State 1 Total Gradient Magnitude (hartrees/Bohr) Difference Contours SA-2/6-311++G(d,p) Level.

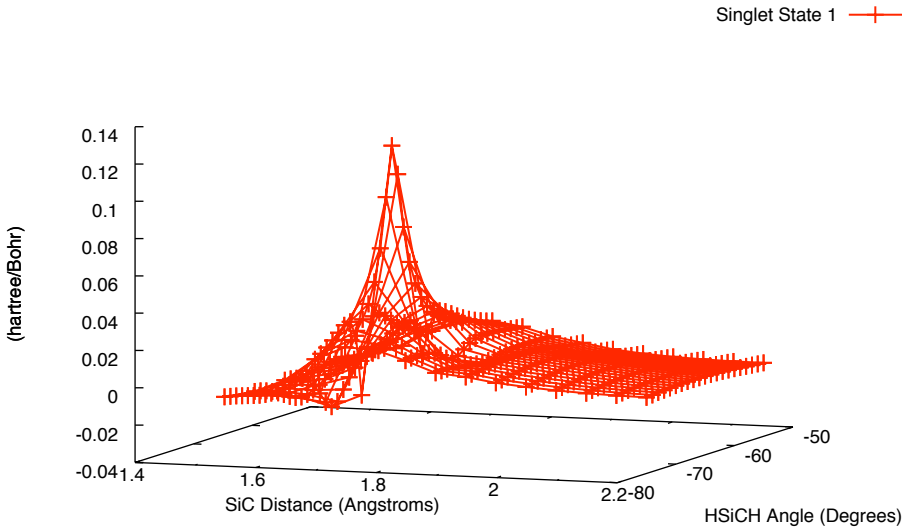


Figure 13. $SiCH_4$ C.I. (ORMAS(scheme 1b) – (8,8)CAS) State 2 Total Gradient Magnitude (hartrees/Bohr) Difference Contours SA-2/6-311++G(d,p) Level.

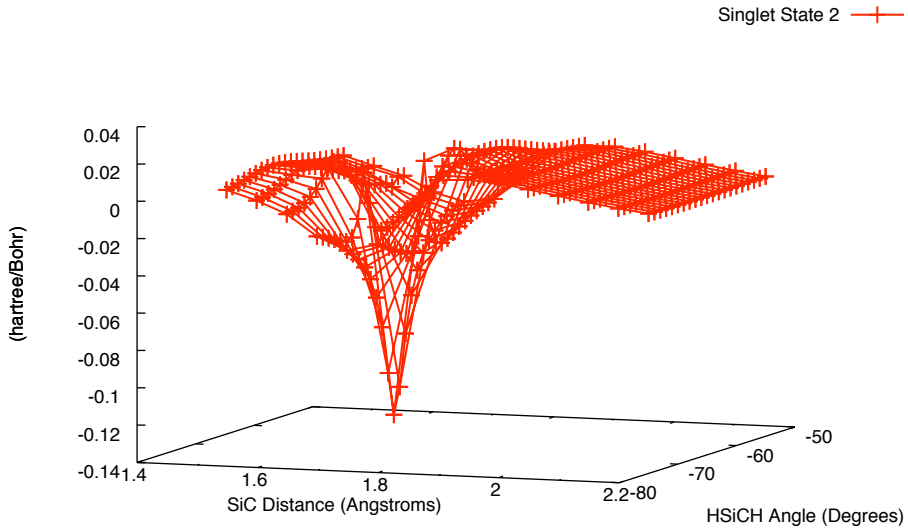


Figure 14. $SiCH_4$ C.I. (ORMAS(scheme 1b) – (8,8)CAS) NACME Normalized Vector Overlap Contours SA-2/6-311++G(d,p) Level.

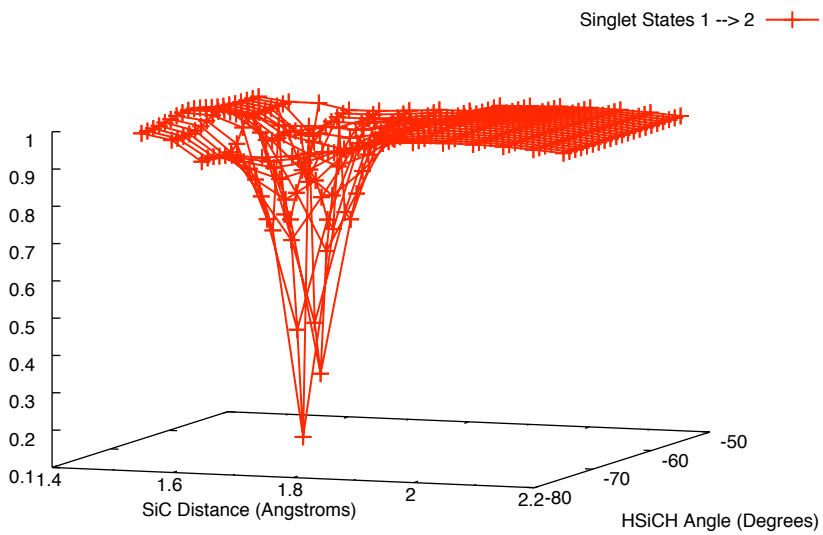


Figure 15. Butadiene C.I. (ORMAS(butadiene1) – (10,10)CAS) Excitation Energy (kcal/mol) Difference Contours SA-2/6-31G(d) Level.

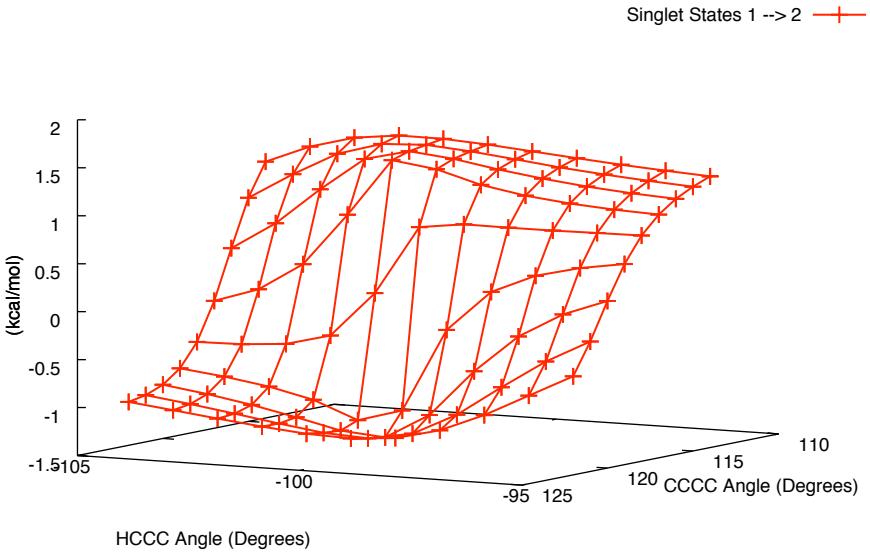


Figure 16. Butadiene C.I. (ORMAS(butadiene1) – (10,10)CAS) Total NACME Magnitude (Bohr^{-1}) Difference Contours SA-2/6-31G(d) Level.

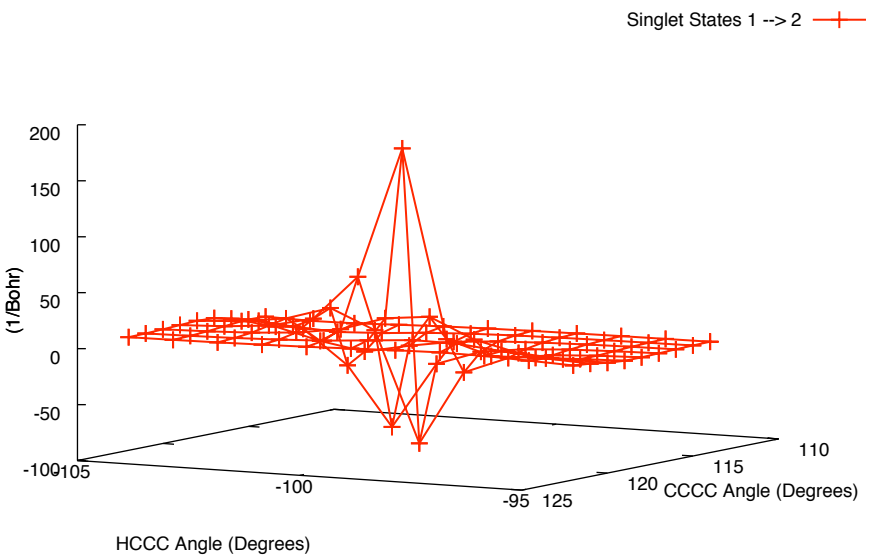


Figure 17. Butadiene C.I. ORMAS(butadiene1) and (10,10)CAS NACME Normalized Vector Overlap Contours SA-2/6-31G(d) Level.

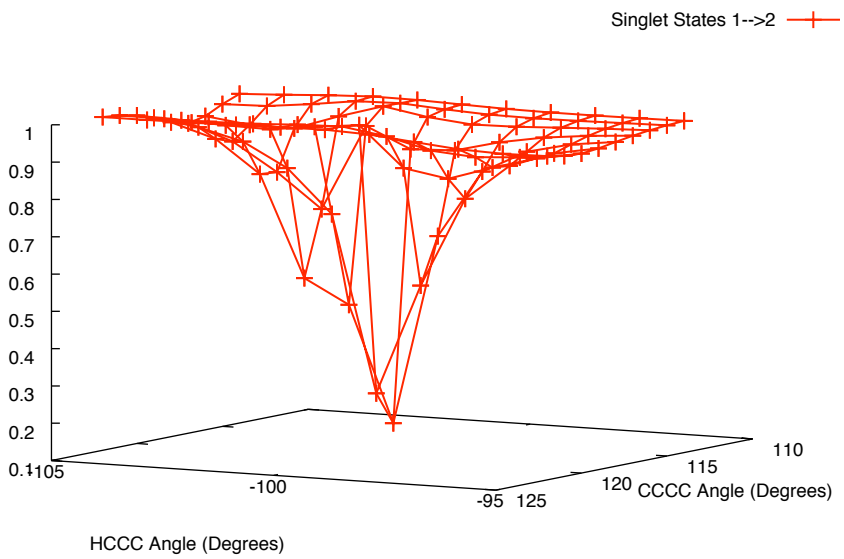


Figure 18. Butadiene C.I. ORMAS(butadiene1) and (10,10)CAS State 1 Gradient Normalized Vector Overlap Contours SA-2/6-31G(d) Level.

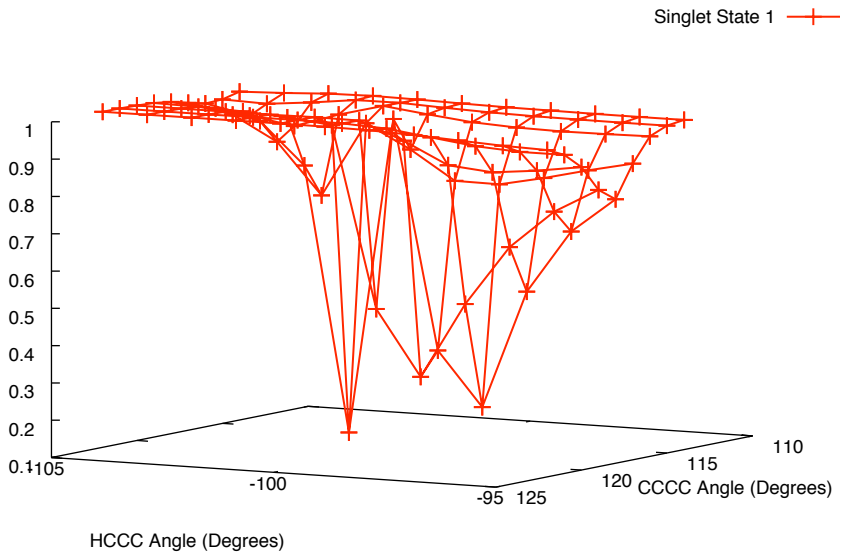


Figure 19. Butadiene C.I. ORMAS(butadiene1) and (10,10)CAS State 2 Gradient Normalized Vector Overlap Contours SA-2/6-31G(d) Level.

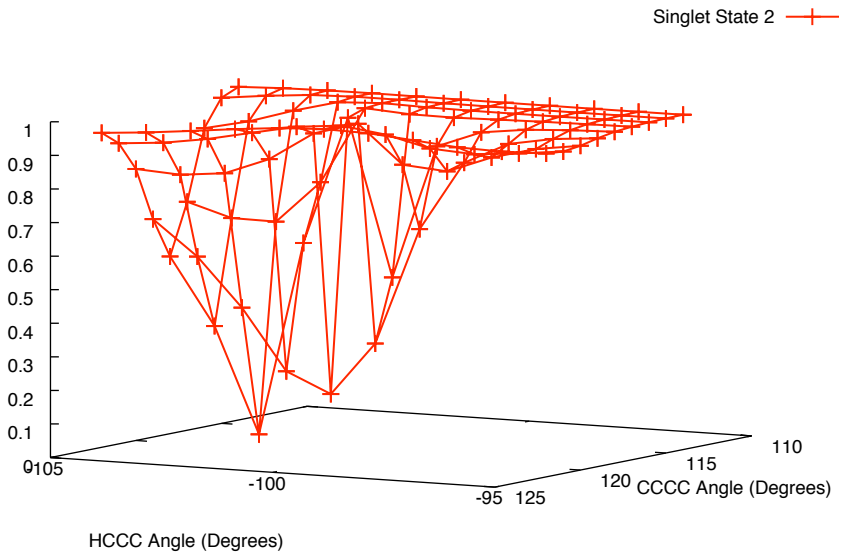
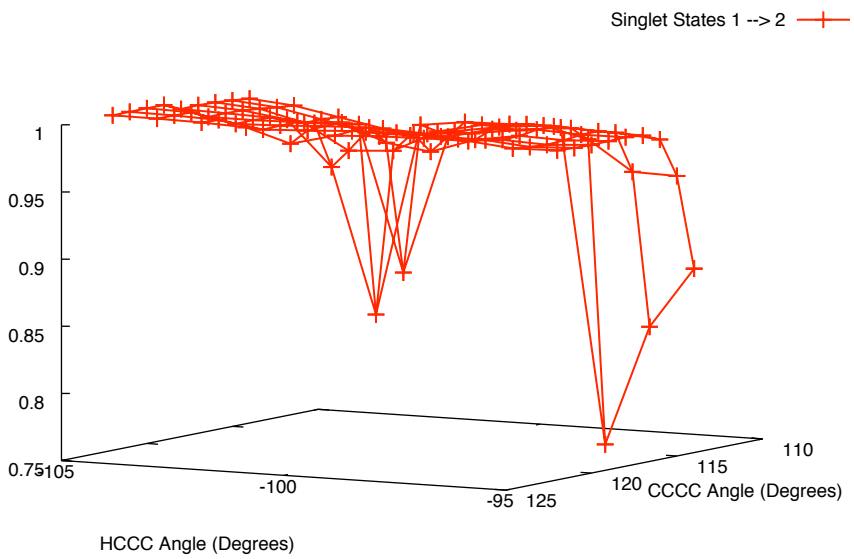


Figure 20. Butadiene C.I. ORMAS(butadiene2) and (10,10)CAS NACME Normalized Vector Overlap Contours SA-2/6-31G(d) Level.



References:

[1]West, A. C.; Dudley, T. J.; Schmidt, M. W.; Gordon, M. S.; Windus, T. L. *unpublished*.

[2]Shepard, R. *Adv. Chem. Phys.* **1987**, *69*, 63.

[3]Ruedenberg, K.; Sundberg, K. R. *Quantum Sci.* **1976**, 505.

[4]Cheung, L. M.; Sundberg, K. R.; Ruedenberg, K. *J. Am. Chem. Soc.* **1978**, *100*, 8024.

[5]Cheung, L. M.; Sundberg, K. R.; Ruedenberg, K. *Int. J. Quantum Chem.* **1979**, *16*, 1103.

[6]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M. *Chem. Phys.* **1982**, *71*, 51.

[7]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 65.

[8]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41.

[9]Feller, D. F.; Schmidt, M. W.; Ruedenberg, K. *J. Am. Chem. Soc.* **1982**, *104*, 960.

[10]Siegbahn, P.; Heiberg, A.; Roos, B.; Levy, B. *Phys. Scr.* **1980**, *21*, 323.

[11]Roos, B. *Adv. Chem. Phys.* **1987**, *69*, 399.

[12] Roos, B. O. *Lect. Notes Chem.* **1992**, *58*, 177.

[13]Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364.

[14]Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9377.

[15]Olsen, J.; Roos, B. O.; Joergensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1988**, *89*, 2185.

[16]Panin, A. I.; Simon, K. V. *Int. J. Quantum Chem.* **1996**, *59*, 471.

[17]Panin, A. I.; Sizova, O. V. *J. Comput. Chem.* **1996**, *17*, 178.

[18]Nakano, H.; Hirao, K. *Chem. Phys. Lett.* **2000**, *317*, 90.

[19]Fleig, T.; Olsen, J.; Marian, C. M. *J. Chem. Phys.* **2001**, *114*, 4775.

[20]Khait, Y. G.; Song, J.; Hoffmann, M. R. *Int. J. Quantum Chem.* **2004**, *99*, 210.

[21]Joergensen, P.; Simons, J. *J. Chem. Phys.* **1983**, *79*, 334.

[22]Brooks, B. R.; Laidig, W. D.; Saxe, P.; Goddard, J. D.; Yamaguchi, Y.; Schaefer, H. F., III. *J. Chem. Phys.* **1980**, *72*, 4652.

[23]Goddard, J. D.; Handy, N. C.; Schaefer, H. F., III. *J. Chem. Phys.* **1979**, *71*, 1525.

[24]Yamaguchi, Y.; Goddard, J. D.; Osamura, Y.; Schaefer, H. F., III *A new Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*, Oxford, New York, 1994.

[25]Tully, J. C. *Mod. Theor. Chem.* **1976**, *2*, 217.

[26]Lengsfield, B. H., III; Yarkony, D. R. *Adv. Chem. Phys.* **1992**, *82*, 1.

[27]Lischka, H.; Dallos, M.; Szalay, P. G.; Yarkony, D. R.; Shepard, R. *J. Chem. Phys.* **2004**, *120*, 7322.

[28]Celani, P.; Werner, H.-J. *J. Chem. Phys.* **2003**, *119*, 5044.

[29]Nakano, H.; Otsuka, N.; Hirao, K. *Recent Adv. Comput. Chem.* **1999**, *4*, 131.

[30]Redmon, L. T. *Phys. Rev. A* **1982**, *25*, 2453.

[31]Khait, Y. G.; Theis, D.; Hoffmann, M. R. *Chem. Phys.* **2012**, *401*, 88.

[32]Farazdel, A.; Dupuis, M. *J. Comput. Chem.* **1991**, *12*, 276.

[33]Yarkony, D. R. *J. Phys. Chem.* **1993**, *97*, 4407.

[34]Manaa, M. R.; Yarkony, D. R. *J. Chem. Phys.* **1993**, *99*, 5251.

[35]Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Mol. Phys.* **2006**, *104*, 1039.

[36]Levine, B. G.; Coe, J. D.; Martinez, T. J. *J. Phys. Chem. B* **2008**, *112*, 405.

[37]Keal, T. W.; Koslowski, A.; Thiel, W. *Theor. Chem. Acc.* **2007**, *118*, 837.

[38]Maeda, S.; Ohno, K.; Morokuma, K. *J. Chem. Theory Comput.* **2010**, *6*, 1538.

[39]Bearpark, M. J.; Robb, M. A.; Schlegel, H. B. *Chem. Phys. Lett.* **1994**, *223*, 269.

[40]Barbatti, M.; Paier, J.; Lischka, H. *J. Chem. Phys.* **2004**, *121*, 11614.

[41]Tao, H.; Levine, B. G.; Martinez, T. J. *J. Phys. Chem. A* **2009**, *113*, 13656.

[42]Pitonak, M.; Lischka, H. *Mol. Phys.* **2005**, *103*, 855.

[43]Zechmann, G.; Barbatti, M.; Lischka, H.; Pittner, J.; Bonacic-Koutecky, V. *Chem. Phys. Lett.* **2006**, *418*, 377.

[44]Olivucci, M.; Ragazos, I. N.; Bernardi, F.; Robb, M. A. *J. Am. Chem. Soc.* **1993**, *115*, 3710.

[45]Krawczyk, R. P.; Malsch, K.; Hohlneicher, G.; Gillen, R. C.; Domcke, W. *Chem. Phys. Lett.* **2000**, *320*, 535.

[46]Dallos, M.; Lischka, H. *Theor. Chem. Acc.* **2004**, *112*, 16.

[47]Ostojic, B.; Domcke, W. *Chem. Phys.* **2001**, *269*, 1.

[48]Gordon, M. S.; Schmidt, M. W. *Theory Appl. Comput. Chem.: First Forty Years* **2005**, 1167.

[49]Docken, K. K.; Hinze, J. *J. Chem. Phys.* **1972**, *57*, 4928.

[50]Ivanic, J.; Ruedenberg, K. *J Comput Chem* **2003**, *24*, 1250.

[51]Lengsfield, B. H., III. *J. Chem. Phys.* **1980**, *73*, 382.

[52]Fletcher, G. D. *Mol. Phys.* **2007**, *105*, 2971.

[53]Yarkony, D. R. *Chem. Phys. Lett.* **1981**, *77*, 634.

[54]Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.

[55]Gordon, M. S. *Chem. Phys. Lett.* **1980**, *76*, 163.

[56]Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

[57]Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

[58]Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

[59]Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.

[60]Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

[61]Dudley, T. J.; Olson, R. M.; Schmidt, M. W.; Gordon, M. S. *J. Comput. Chem.* **2006**, *27*, 352.

CHAPTER 6. NEWTON-X-GAMESS INTERFACE: DYNAMICS APPLICATION TO METHANIMINIUM WITH ORMAS

Aaron C. West, Mario Barbatti, and Theresa L. Windus

Abstract:

The program package NEWTON-X and the electronic structure suite GAMESS are interfaced for nonadiabatic and adiabatic "on-the-fly" dynamics simulations. Using GAMESS with NEWTON-X expands on the wavefunctions available for use in dynamics studies. In particular, this interface allows for the first study of nonadiabatic dynamics with the ORMAS approximation, which is unique to GAMESS. The methaniminium cation is used for benchmark calculations at the CASSCF and ORMAS levels of theory. For the CASSCF and ORMAS runs, several dynamics simulations were performed with various computationally feasible active space choices (or schemes) in order to test the qualitative accuracy and relative expense of different active space choices. Overall, for ORMAS orbital subspace divisions, schemes with no excitations between orbital subspaces give qualitatively incorrect state populations while schemes with single excitations between orbital subspaces recover the qualitatively correct state populations relative to the CASSCF level of theory.

Introduction:

The NEWTON-X[1] program provides dynamical reaction rates for ultrafast processes as found for example in photochemistry. In order to avoid the more expensive so-called two-step approach, the potential energy surface (PES) is not mapped out before dynamics simulations but rather generated "on-the-fly" by a 'swarm' of independent trajectories. In

this manner, NEWTON-X initially generates a UV absorption spectrum and subsequently performs nonadiabatic or adiabatic dynamics as appropriate. The simulated UV absorption spectrum derives from the transition probabilities of a set of initial conditions (ICs), which consist of a sampling of coordinates and momenta. First, the distribution of the coordinates and momenta is taken from a given harmonic vibrational state. Second, from this distribution oscillator strengths and energy differences between electronic states are computed and ideally simulate a "Franck-Condon" wavepacket for an ultra short pulse. Various restrictions may be imposed (i.e. energy restrictions for initiating a trajectory). For nonadiabatic dynamics simulations, the nuclear motions are based on classical trajectories where the '1990' Tully surface hopping (TSH)[2] method defines the instantaneous nuclear forces. Based on the nuclear positions and a chosen electronic level of theory, an electronic structure suite then provides an approximate wavefunction from which the next set of nuclear forces are generated. Along each trajectory, the NEWTON-X program repeatedly executes an electronic structure calculation by calls through a shell script. This loose coupling of the two programs almost effortlessly allows ongoing development of both NEWTON-X and the external electronic structure programs with little change to the interface. In addition, the electronic structure calculations are by far the most expensive computations relative to time step integrations along each trajectory in the surface hopping methods, so additional overhead by calling the script is negligible. NEWTON-X is interfaced with several third party programs, which include COLUMBUS,[3-5] TURBOMOLE,[6] DFTB,[7] TINKER,[8] GAUSSIAN,[9] MOPAC,[10,11] ACES II,[12] and DFT-MRCI[13]. The interface with GAMESS[14] will now provide additional wavefunction choices for both nonadiabatic and adiabatic dynamics.

In this work, both adiabatic and nonadiabatic dynamics runs are implemented for the NEWTON-X-GAMESS interface. Many levels of theory (e.g. MP2, CCSD(T)) are available for adiabatic dynamics while nonadiabatic dynamics in GAMESS are currently limited to multiconfigurational self-consistent field (MCSCF)[15,16] implementations. Adiabatic runs require an energy and gradient for the electronic state of interest while nonadiabatic runs require energies and nuclear derivatives for several electronic states. The present paper demonstrates the NEWTON-X-GAMESS interface with benchmark dynamics calculations on the methaniminium cation CNH_4^+ at the complete active space self-consistent field (CASSCF)[15-20] level of theory. Also, in conjunction with a previous study,[21] this work examines nonadiabatic results at the occupation restricted multiple active space (ORMAS)[22,23] level. The ORMAS approximation restricts excitations between different invariant orbital subspaces in order to reduce the overall configuration interaction (CI) expansion size in an *a priori* fashion. Partitioning a total active space into several subspaces and defining maximum and minimum electron occupation values for each subspace manifests the excitation restriction of electrons between the subspaces. So, in addition to the NEWTON-X-GAMESS interface, this work provides the first initial investigation with ORMAS wavefunctions in nonadiabatic dynamics simulations.

Methods:

Figure 1 gives the overall processes used to generate the ICs, trajectories, and statistical results. Geometry optimizations, Hessians, energies, and oscillator strengths are calculated by an electronic structure code and are used to produce the ICs, which in turn generate the starting coordinates and momenta for the trajectories. For each trajectory at each time step, energies, oscillator strengths, gradients, and nonadiabatic couplings (NACs)

are generated by the electronic structure code and used to produce new nuclear velocities and coordinates.

The loose coupling of the NEWTON-X-GAMESS interface uses Perl scripts and disk-based files to complete the simulations. The GAMESS user creates a typical NAC input, which serves as a template for GAMESS executions during each trajectory. In addition, NEWTON-X input files are required to specify the dynamics simulation parameters. NEWTON-X reads the GAMESS input file, makes modifications based on initial or updated coordinates, and then writes a new GAMESS input file to be executed. The Perl scripts in NEWTON-X then launch a GAMESS job to obtain energies, gradients, NACs, and oscillator strengths. Once this job is complete, NEWTON-X extracts the quantities from the GAMESS output files for the next step in the simulation. NEWTON-X also makes determinations on whether a trajectory should be discontinued due to energy conservation problems or other errors.

To test this implementation and examine the use of ORMAS for dynamics calculations, nonadiabatic simulations of the methanimium cation are performed. All calculations were performed with NEWTON-X[1] and GAMESS[14] software in serial on a 2 dual-core AMD Opteron processor with 8 GB of RAM. State averaged[24] (SA)- and single-state (i.e. non-SA) MCSCF calculations were carried out at the CASSCF[15-20] and ORMAS[22,23] levels of theory for the lowest-lying singlet states. For the SA calculations, the three lowest-lying singlet states were SA with equal weights. The determinant-based method and full Newton-Raphson converger with augmented Hessian technique were utilized.[25-27] Because previous studies[1,28] at the SA-3-(6, 4)CAS/aug-cc-pVDZ and SA-3-(4,3)CAS/6–31G(d) levels already exist for comparisons, the 6-31G(d)[29-31] basis set was used for all

calculations. In addition, the previous studies demonstrate that for various static calculations, the CASSCF level of theory will recover qualitative results relative to the multireference configuration-interaction (MR-CI) level.[1,28] Based on computational feasibility and on orbitals already known[1] to be important to configurations in CNH_4^+, several active spaces were chosen for the CASSCF and ORMAS levels (see Results Section). In addition to these separate levels of theory, (12,8)CAS ICs were generated for initial use in dynamics runs at chosen ORMAS levels of theory to examine the IC impact on the simulations but made no difference to the results (see Supporting Information).

At each single-state level of theory, minimum geometry optimizations and Hessians were performed on the lowest-lying singlet. At each level of theory but with SA-MCSCF, 100 ICs and trajectories were run. The uncorrelated quantum harmonic oscillator (Wigner)[32] distribution was used to generate the ICs distribution without excitation energy restrictions. Each trajectory was initiated on the 2nd excited singlet state and allowed a 100 fs maximum simulation time with 0.5 fs time steps. The time-dependent Schrodinger equation was integrated with the 5th order Butcher integrator[33] with 0.025 fs time intervals (i.e. sub-time steps), and a 1st order decoherence correction[34] was applied to the time-dependent coefficients. Then, the Hammes-Schiffer and Tully hopping algorithm[35] was used to generate effective transition probabilities and thus select the instantaneous adiabatic force in all trajectories. A random seed was selected for random number generation. In addition, after a frustrated hop, the momentum direction was kept, and after an inter-state hop, the momentum was adjusted in the direction of the nonadiabatic coupling vector[36-38].

Because of the difficulties involved in converging single excitations at the ORMAS level, some tolerances were adjusted. In the ORMAS(scheme2) trajectory runs (the schemes

are described below), the default GAMESS Lagrangian asymmetry (energy difference) tolerance was increased from $1x10^{-5}$ ($1x10^{-10}$) to $1x10^{-4}$ ($1x10^{-8}$); these only slightly less stringent tolerances are the default tolerances in COLUMBUS for nonadibatic dynamics.[3-5] Also, because of longer and impractical convergence times in the analytical derivatives in the ORMAS(scheme3) and ORMAS(scheme4) trajectory runs, the GAMESS MCSCF analytical gradient and NAC tolerances were increased from $1x10^{-6}$ to $1x10^{-4}$.

Results and Discussion:

This study uses several MCSCF levels of theory for comparisons. Table 1 contains the active orbitals for each MCSCF level of theory along with the definition of each active orbital subspace. Table 1 also defines the x(2,1) and x(2,2) notations where x gives the total number of each subspace. The following MCSCF schemes are presented: (12,8)CAS, (4,4)CAS, 2(2,2)ORMAS with no excitations between spaces (ORMAS(scheme1)), 2(2,2)ORMAS with single excitations between spaces (ORMAS(scheme2)), 4(2,1)+2(2,2)ORMAS with no excitations between spaces (ORMAS(scheme3)), and 4(2,2)+2(2,2)ORMAS with no excitations between spaces (ORMAS(scheme4)). The active orbital subspaces were chosen based on two major criteria. As previously shown[1] for the three lowest-lying singlet states, the most important configurations result from the CN σ,σ*,π, and π* orbitals. For the (4,4)CAS for the lowest-lying (S_0), first-excited (S_1), and second-excited (S_2) singlet states, the most dominant configurations at the planar geometry are CN πσ, CN πσ*, and CN ππ*, respectively. An approximate MCSCF wavefunction should include at least these configurations (or associated determinants for a determinantal implementation) in order to be qualitatively appropriate. So, based on this criterion (criterion 1), all the active spaces contain at least the CN σ,σ*,π, and π* orbitals. In addition, as will be

shown, larger MCSCF active spaces that include heavy element and hydrogen σ bonds (or rather smaller MCSCF core spaces that include only the chemical core orbitals) can result in longer simulation times for certain trajectories (i.e. trajectory survival). The NEWTON-X program terminates any trajectory that violates total energy conservation, which can occur if a different orbital rotates into the active space and results in an energy change above a certain tolerance. So, based on this criterion (criterion 2), larger active spaces (e.g. ORMAS(scheme3) in Table 1) will facilitate additional comparisons.

In terms of immediate limitations on the MCSCF orbital spaces, any attempts to make a "4(2,1)+(4,4)" or "4(2,2)+(4,4)" ORMAS calculation without single excitations resulted in an improper active space for SA wavefunctions at the optimized geometries. As a result of these failures and single excitation convergence issues, the "4(2,1)+2(2,2)" or "4(2,2)+2(2,2)" calculations with single excitations were not attempted.

In all cases, the geometry optimized to the same planar structures (see Figure 2) for the lowest-lying singlet state. Table 2 gives a summary of various statistics for each dynamics simulation that include: number of trajectories terminations due to energy conservation violation (i.e. NEWTON-X energy error trajectory count), number of trajectories that reach a MCSCF excessive iteration error (which occurs as certain geometry distortions lead to active orbital instabilities that prevent MCSCF energy convergence) (i.e. MCSCF excessive iteration error), number of trajectories with an MCSCF states error (which occurs from solving for too few or too many CI states) (i.e. MCSCF states error), number of trajectories that actually reach 100 fs (i.e. 100 fs trajectory count), number of trajectories with at least one state-state energy difference of less than ~4.6 kcal/mol for all SA states (i.e. possible conical intersection (C.I.) trajectory count), number of time steps with state-state

energy differences of less than ~4.6 kcal/mol for all SA states (i.e. time steps with possible C.I.s for state $S_0 \rightarrow S_1$ and state $S_1 \rightarrow S_2$), and number of trajectories that lack conservation of probability and reach a hopping probability larger than 1 (i.e. NEWTON-X >1 probability). In addition, no trajectories error out from NAC convergence issues. As shown in Table 2, most simulation errors resulted from NEWTON-X energy conservation issues between steps. These simulation errors usually occurred when different orbitals rotated into the active space.

In order to show the expense of the different schemes, Table 3 gives the number of determinants and average computational run times for selected CN dissociating trajectories. Overall, Table 3 shows that computational expense quickly becomes unmanageable for any scheme larger than ORMAS(scheme4). For a single run (i.e. energy, gradients, and NACs), the ORMAS(scheme4) (corresponding (12,12)CAS) has 13,236 (853,776) determinants and a computational run time of ~16 minutes (~2.75 days). This example shows how the ORMAS approximation drastically reduces timings relative to CASSCF runs with comparable numbers of active orbitals.

Because the average adiabatic populations closely match the fractions of trajectories for each electronic state, the figures show only fractions. For the (12,8)CAS results, three figures (Figures 3 – 5) are given for benchmarking and comparison to schemes with smaller active spaces. Because of the different levels of theory for both the ICs and the trajectories, these figures differ slightly from figures given in the previous benchmark study;[1] however, no major qualitative differences are found between the results of the two different studies. Figure 3 shows the cross-section as a function of excitation energy from the ICs. Figure 4 shows the S_2, S_1, and S_0 population versus time over the 100 trajectories. As will be shown below, the depopulation of the S_2 state within about the first 10 fs is of particular interest.

Figure 5 shows the CN distance versus time for 100 trajectories. The (12,8)CAS model has only two orbitals in the MCSCF core – the C1s and N1s chemical core orbitals – and does not include the CH or NH σ* orbitals. So, the plotted trajectories in Figure 5 are quite 'thick' (i.e. most trajectories do not get terminated). This 'thickness' is lost in many of the active space schemes discussed below, and this loss becomes even more severe in schemes where the MCSCF core contains more than just N1s and C1s orbitals.

The (4,4)CAS ICs spectrum and state populations graphs resemble the corresponding (12,8)CAS plots. So, these figures are omitted; any figures omitted can be found in the Supporting Information. However, the (4,4)CAS trajectories plot does not display the same 'thickness' after 70 fs as the corresponding (12,8)CAS plot, so these results are not as statistically valid. (Since similar trajectories graphs are given for other small active space schemes below, this graph is also omitted.) The (4,4)CAS state populations graph (omitted) also shows this loss of statistical validity through the comparisons of the S_2 and S_1 state populations after 70 fs. However, the large population transfer before 10 fs between the S_2 and S_1 states is still present in the (4,4)CAS.

For the ORMAS(scheme1) results, three figures are given. The ICs spectrum (Figure 6) shows smaller cross-sections with approximately the same excitation energies; however, using (12,8)CAS ICs to initiate the ORMAS(scheme1) dynamics produces identical results with only slightly more S_2 depopulation. The state populations (Figure 7) do not show the S_2 - S_1 population transfer in the first 10 fs. In addition, plotting the trajectories (Figure 8) demonstrates the impact of NEWTON-X terminations from energy discontinuities (Table 2) on smaller active space schemes. However, as shown in Figure 7, some population transfer occurs at later times in the simulation. In addition, in order to allow for a direct comparison

with ORMAS(scheme2) with the larger MCSCF tolerances, ORMAS(scheme1) was also run with these larger MCSCF tolerances. These additional ORMAS(scheme1) trajectory calculations gave identical results.

The ORMAS(scheme2) ICs spectrum and trajectories plots resemble the corresponding graphs from the (4,4)CAS and thus are omitted. The ORMAS(scheme2) state populations (Figure 9) strongly resemble the (12,8)CAS (Figure 4) and (4,4)CAS state populations. Thus, for the CNH_4^+ system as in our previous stationary point work[21], the single excitations are *essential* to obtaining qualitatively correct results.

The ORMAS(scheme3) ICs spectrum resembles the ORMAS(scheme1) ICs spectrum and thus is omitted. With no excitations between active orbital subspaces, the state populations (Figure 10) also lacks the qualitative S_2 - S_1 population transfer in the first 10 fs. The ORMAS(scheme3) trajectories graph (omitted) looks similar to Figure 6 at the ORMAS(scheme1) level but has more 'thickness' from a larger active space. As a result, the S_2 state is ~20% more depopulated than at the ORMAS(scheme1) level. In part, this effect demonstrates the importance of active space stability within a given number of trajectories. However, initiating ORMAS(scheme1) with (12,8)CAS also gives more S_2 depopulation. So, these results are very slightly sensitive to the ICs.

The ORMAS(scheme4) ICs spectrum (omitted) does not resemble the other spectra. However, two sets of ORMAS(scheme4) trajectories were initiated with ICs at the ORMAS(scheme4) and (12,8)CAS levels. The results from these two simulations are identical. The state populations and trajectories graphs resemble the corresponding plots at the ORMAS(scheme3) level of theory up to about 70 fs and thus are also omitted. As shown and explained in Table 2, this particular scheme shows problems for the MCSCF states error.

Conclusions:

The NEWTON-X-GAMESS interface is implemented and tested with the methaniminium cation, CNH_4^+, benchmark system. For this benchmark, the defaults in NEWTON-X were used, and (12,8)CAS results reproduce the characteristic 10 fs S_2 depopulation relative to previous studies[1,28] on CNH_4^+. In addition, ICs and trajectories were run for several active space schemes at the CASSCF and ORMAS levels of theory. With respect to longer trajectory simulation times, the most successful runs contain only the C1s and N1s orbitals in the MCSCF core and experience less energy conservation issues. However, these runs are expensive. MCSCF methods with smaller active spaces and additional limited excitations into the external space might alleviate some of the energy conservation issues if computational expense does not pose an issue. For the CNH_4^+ test case, the ORMAS approximation with no excitations between orbital subspaces does not reproduce the characteristic 10 fs S_2 depopulation. However, the single excitation ORMAS approximation does produce this depopulation. In addition, different ORMAS schemes might help to identify different state population transfers with individual excitation schemes. In this sense, following different trajectory types with different ORMAS schemes might prove useful in more complicated cases.

Acknowledgments:

ACW and TLW are indebted to Michael W. Schmidt and Mark S. Gordon for help in using the capabilities of GAMESS and MCSCF. ACW would like to thank both Mario Barbatti and TLW for their patience and help with the interface process. This material is based upon work supported by the National Science Foundation under Grant No. OISE-0730114 for the Partnerships in International Research and Education (PIRE). TLW

acknowledges computing resources purchased through funds provided by Ames Laboratory and Iowa State University.

Supporting Information Available:

Supporting Information contains additional tables and Figures that were omitted.

Table 1. Active orbitals are given for each MCSCF level of theory.

MCSCF level of theory	Active orbitals
(12,8)CAS	CNσ, CNσ*, CNπ, CNπ*, 2 CHσ, 2NHσ
(4,4)CAS	CNσ, CNσ*, CNπ, CNπ*
2(2,2)ORMAS no excitations i.e. ORMAS(scheme1)	Space 1 = CNσ, CNσ* Space 2 = CNπ, CNπ*
2(2,2)ORMAS single excitations i.e. ORMAS(scheme2)	Space 1 = CNσ, CNσ* Space 2 = CNπ, CNπ*
4(2,1)+2(2,2)ORMAS no excitations i.e. ORMAS(scheme3)	Space 1 = CHσ Space 2 = CHσ Space 3 = NHσ Space 4 = NHσ Space 5 = CNσ, CNσ* Space 6 = CNπ, CNπ*
4(2,2)+2(2,2)ORMAS no excitations i.e. ORMAS(scheme4)	Space 1 = CHσ,σ* Space 2 = CHσ,σ* Space 3 = NHσ,σ* Space 4 = NHσ,σ* Space 5 = CNσ, CNσ* Space 6 = CNπ, CNπ*

Table 2. Trajectory and time step information counts for various items are given for each MCSCF level of theory. (NX stands for NEWTON-X in this table.)

Counts[a]	(12,8)CAS	(4,4)CAS	ORMAS (scheme1)	ORMAS (scheme2)	ORMAS (scheme3)	ORMAS (scheme4)
NX energy error trajectory	23	57	89	70	80	58
MCSCF excessive iteration error	13	21	2	5	2	2
MCSCF states error[b]	0	0	0	0	0	35
100 fs trajectory	62	22	2	14	17	2
Possible C.I. trajectory	94	69	27	69	46	28

Counts[a]	(12,8)CAS	(4,4)CAS	ORMAS (scheme1)	ORMAS (scheme2)	ORMAS (scheme3)	ORMAS (scheme4)
Time steps with possible C.I. $S_0 \rightarrow S_1$	485	343	188	136	388	81
Time steps with possible C.I. $S_1 \rightarrow S_2$	769	458	35	357	91	38
NX >1 probability	1	0	6	5	1	3

a. Due to other various errors, trajectory counts do not add up to 100.

b. GAMESS number of states error occurs as follows. 1) Too few roots are requested, and the spin states of interest are not captured. 2) Too many roots are requested, and the spins do not converge. This problem results because in a determinant-based implementation without the diagonalization of an individual sub-block in the spin-squared matrix, the Davidson converger mixes up states with equal energies in spite of different spins.

Table 3. Run times in seconds for selected trajectories are given at each MCSCF level of theory. Every selected trajectory shows some CN dissociation in order to obtain reasonable statistics for the average times.

Sample trajectory	(12,8)CAS	(4,4)CAS	ORMAS (scheme1)	ORMAS (scheme2)	ORMAS (scheme3)	ORMAS (scheme4)[a]
Average computer time (s) per 0.5 fs time step	79.5	12.6	11.4	14.0	61.9	955.0[b]
Number of determinants	784	36	18	34	18	13,236[c]

a. For comparison, a (12,12)CAS was performed from the same initial data (i.e. geometry, starting orbitals, etc). One (12,12)CAS energy, gradients, and NAC single calculation finished after ~2.75 days.

b. Because trajectories are not fully dissociated, these timings are probably underestimated.

c. For comparison, the number of determinants for (12,12)CAS is 853,776.

Figure 1. Overall processes for generating initial conditions and trajectories are given.

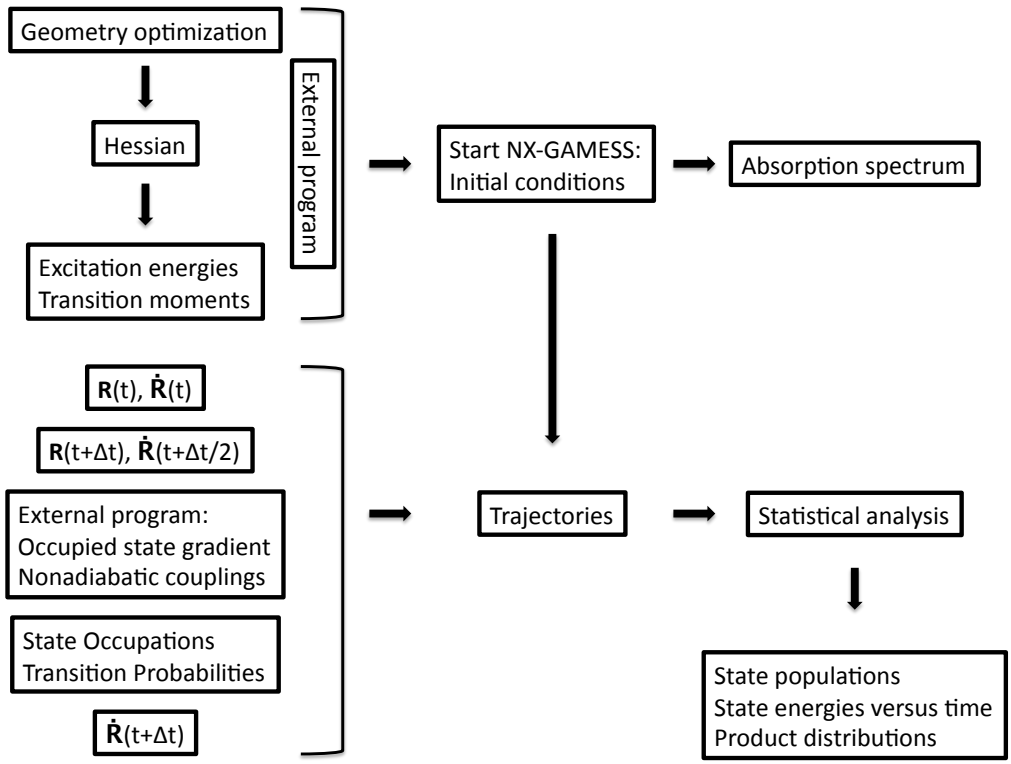


Figure 2. The typical planar structure for the CNH_4^+ lowest-lying singlet minimum is found with geometry optimization.

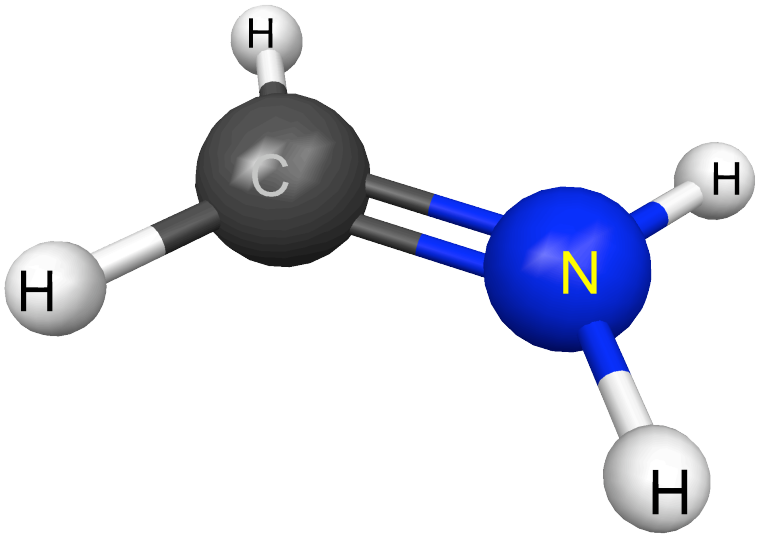


Figure 3. Cross-section ($Å^2$ $molecule^{-1}$) as a function of excitation energy (eV) from ICs is shown at the SA-3-(12,8)CAS level of theory.

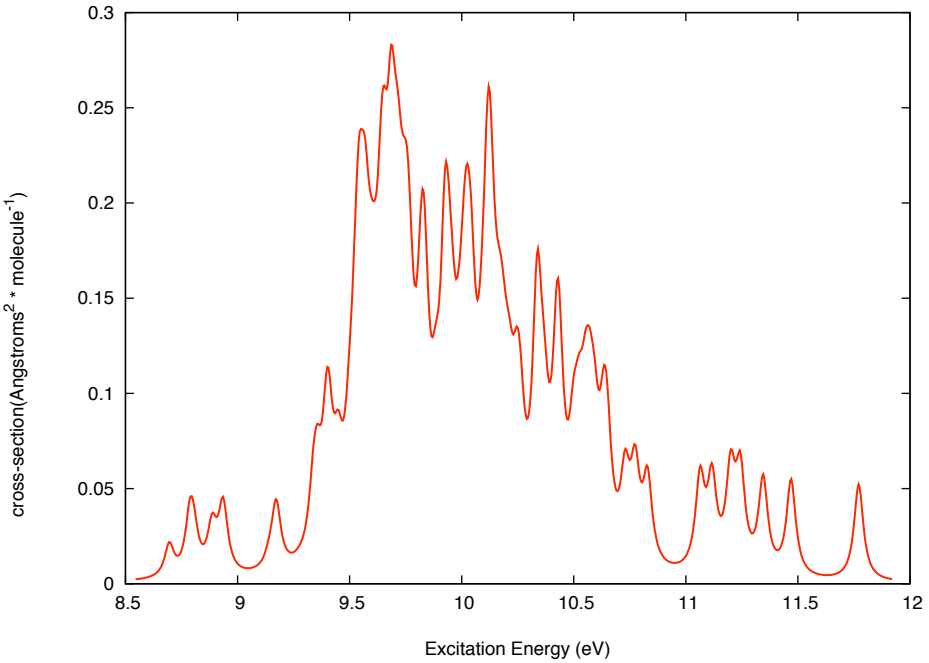


Figure 4. S_2, S_1, and S_0 population (fraction of trajectories and average adiabatic) versus time (fs) is shown at the SA-3-(12,8)CAS level of theory.

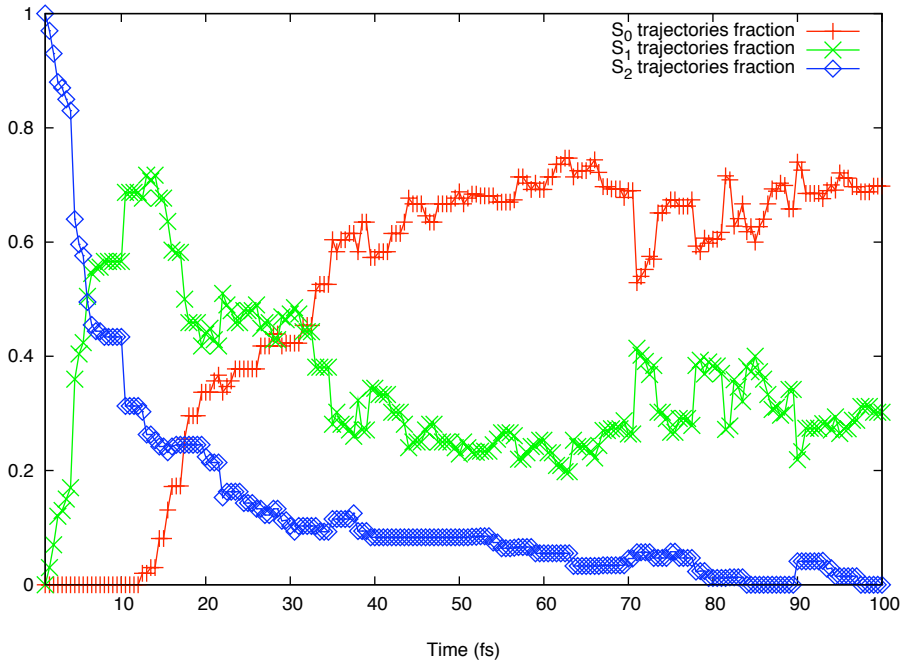


Figure 5. The CN distance (Å) versus time (fs) is shown for 100 trajectories at the SA-3-(12,8)CAS level of theory.

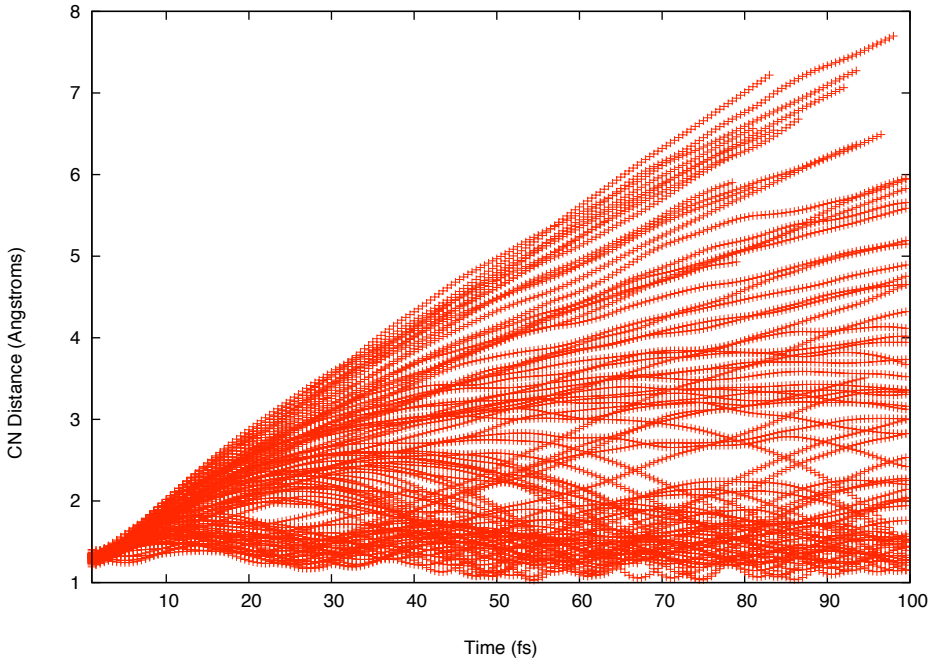


Figure 6. Cross-section ($Å^2$ $molecule^{-1}$) as a function of excitation energy (eV) from ICs is shown at the SA-3-ORMAS(scheme1) level of theory.

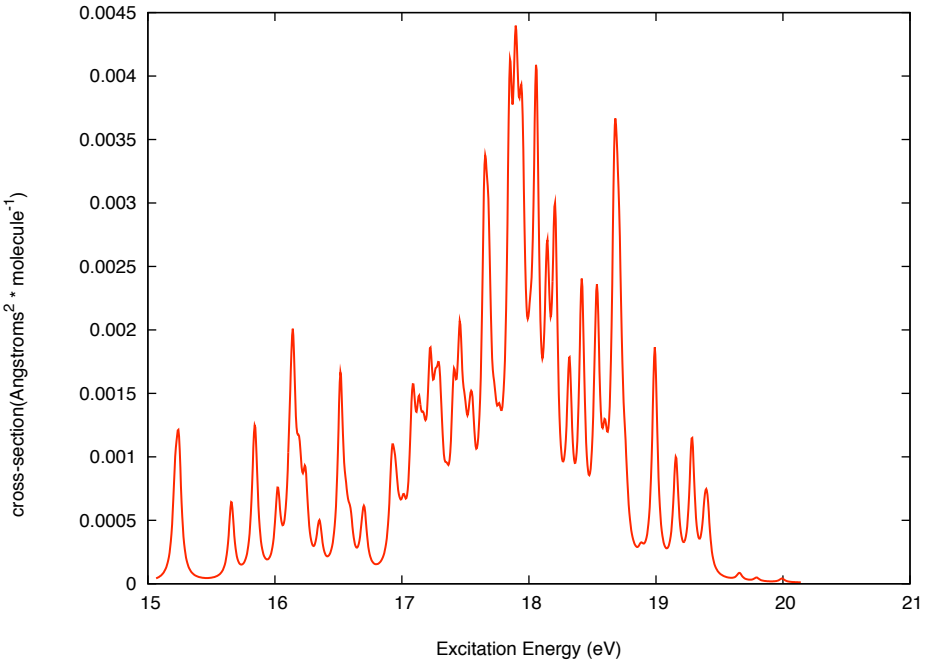


Figure 7. S_2, S_1, and S_0 population (fraction of trajectories and average adiabatic) versus time (fs) is shown at the SA-3-ORMAS(scheme1) level of theory.

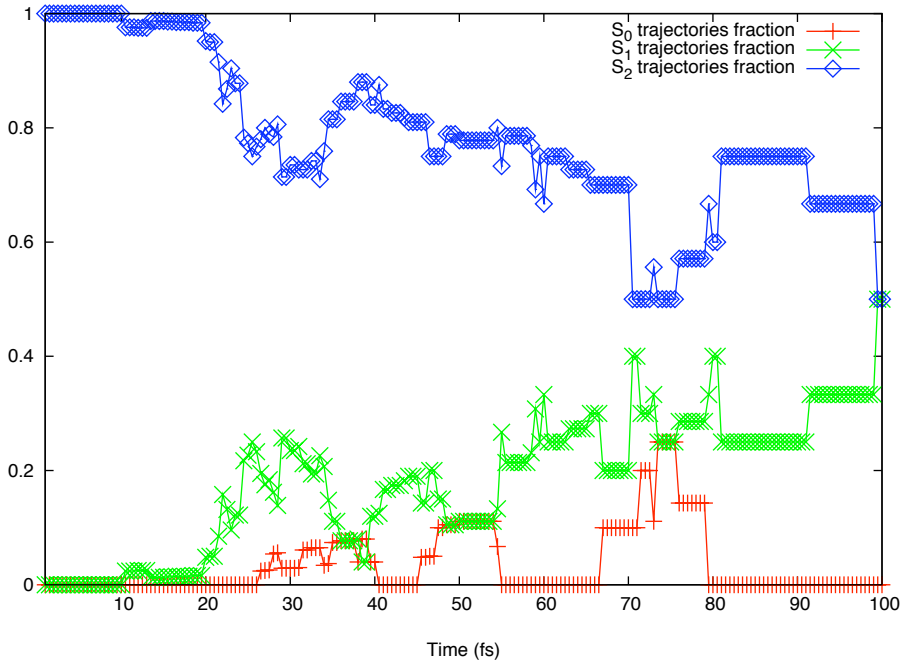


Figure 8. The CN distance (Å) versus time (fs) is shown for 100 trajectories at the SA-3-ORMAS(scheme1) level of theory.

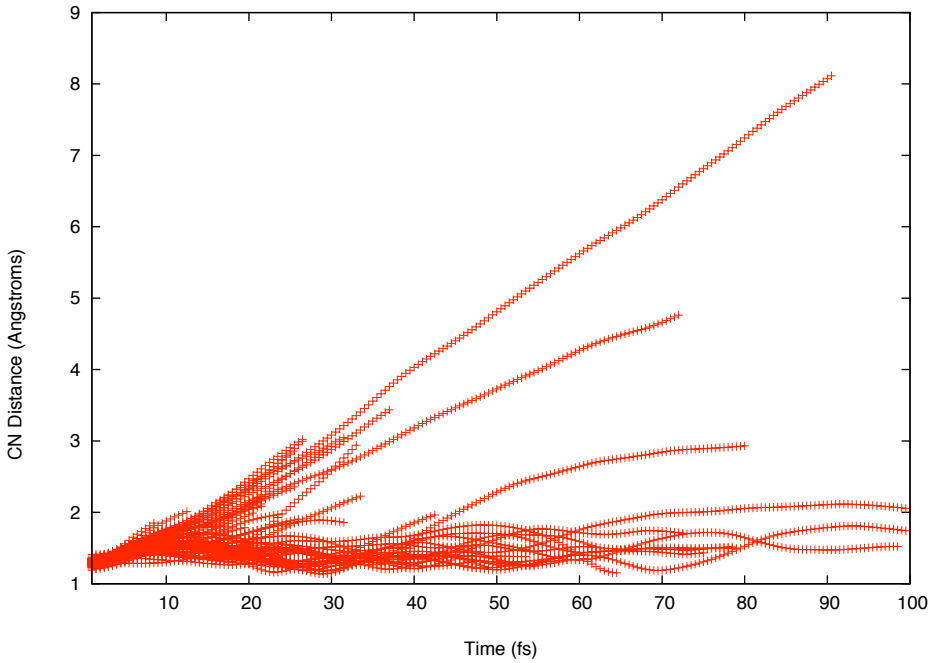


Figure 9. S_2, S_1, and S_0 population (fraction of trajectories and average adiabatic) versus time (fs) is shown at the SA-3-ORMAS(scheme2) level of theory.

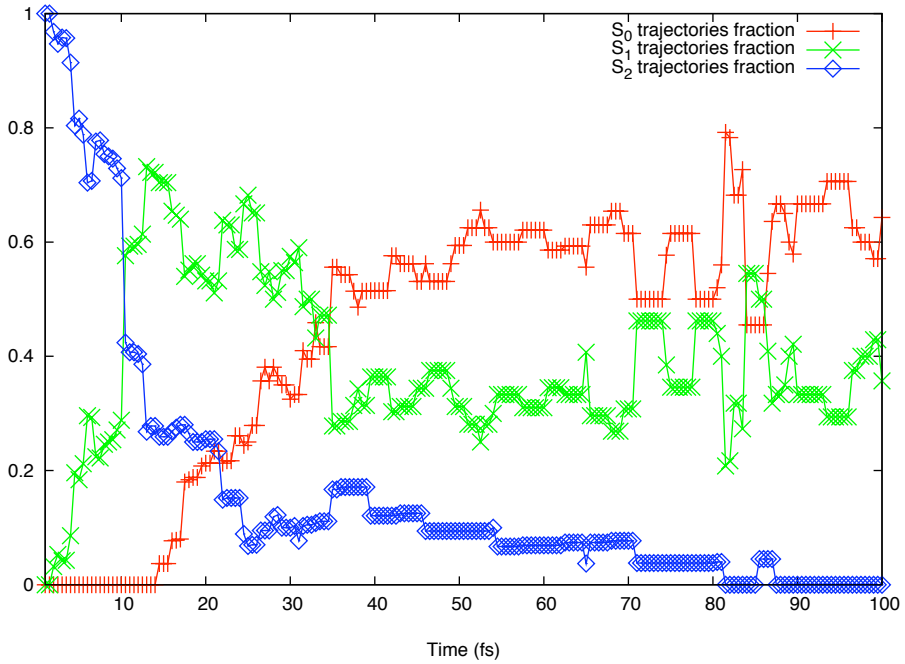
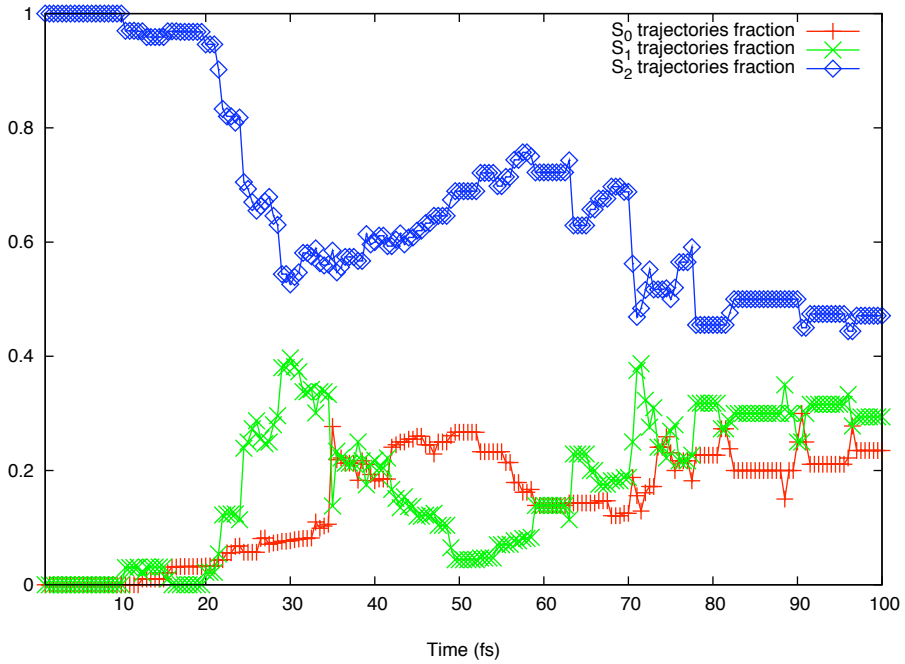


Figure 10. S_2, S_1, and S_0 population (fraction of trajectories and average adiabatic) versus time (fs) is shown at the SA-3-ORMAS(scheme3) level of theory.



References:

[1]Barbatti, M.; Granucci, G.; Persico, M.; Ruckenbauer, M.; Vazdar, M.; Eckert-Maksic, M.; Lischka, H. *J. Photochem. Photobiol., A* **2007**, *190*, 228.

[2]Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061.

[3]Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Muller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664.

[4]Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Mueller, T.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Boehm, H. J.; Chang, A.; Comeau, D. C.; Gdanitz, R.; Dachsel, H.; Ehrhardt, C.; Ernzerhof, M.; Hoechtl, P.; Irle, S.; Kedziora, G.; Kovar, T.; Parasuk, V.; Pepper, M. J. M.; Scharf, P.; Schiffer, H.; Schindler, M.; Schueler, M.; Seth, M.; Stahlberg, E. A.; Zhao, J.-G.; Yabushita, S.; Zhang, Z.; Barbatti, M.; Matsika, S.; Schuurmann, M.; Yarkony, D. R.; Brozell, S. R.; Beck, E. V.; Blaudeau, J.-P. **2006**, www.univie.ac.at/columbus.

[5]Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Mueller, T.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Boehm, H. J.; Chang, A.; Comeau, D. C.; Gdanitz, R.; Dachsel, H.; Ehrhardt, C.; Ernzerhof, M.; Hoechtl, P.; Irle, S.; Kedziora, G.; Kovar, T.; Parasuk, V.; Pepper, M. J. M.; Scharf, P.; Schiffer, H.; Schindler, M.; Schueler, M.; Seth, M.; Stahlberg, E. A.; Zhao, J.-G.; Yabushita, S.; Zhang, Z.; Barbatti, M.; Matsika, S.; Schuurmann, M.; Yarkony, D. R.; Brozell, S. R.; Beck, E. V.; Blaudeau, J.-P.; Ruckenbauer, M.; Sellner, B.; Plasser, F.; Szymczak, J. J. **2008**, www.univie.ac.at/columbus.

[6]Ahlrichs, R.; Baer, M.; Haeser, M.; Horn, H.; Koelmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.

[7]Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316.

[8]Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016.

[9]Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J.; J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. **2004**.

[10]Stewart, J. J. "MOPAC manual. Sixth edition. A general molecular orbital package," Frank J. Seiler Res. Lab., U.S. Air Force Acad., CO, USA., 1990.

[11]Stewart, J. J. P. *MOPAC 2000 and MOPAC 2002. Fujitsu Limited, Tokio, Japan.*

[12]Stanton, J. F.; Gauss, J.; Watts, J. D.; Lauderdale, W. J.; Bartlett, R. J. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1992**, *26*, 879.

[13]Grimme, S.; Waletzke, M. *J. Chem. Phys.* **1999**, *111*, 5645.

[14]Gordon, M. S.; Schmidt, M. W. *Theory Appl. Comput. Chem.: First Forty Years* **2005**, 1167.

[15]Ruedenberg, K.; Sundberg, K. R. *Quantum Sci* **1976**, 505.

[16]Siegbahn, P.; Heiberg, A.; Roos, B.; Levy, B. *Phys. Scr.* **1980**, *21*, 323.

[17]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M. *Chem. Phys.* **1982**, *71*, 51.

[18]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 65.

[19]Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41.

[20]Roos, B. *Adv. Chem. Phys.* **1987**, *69*, 399.

[21]West, A. C.; Windus, T. L. *Theor. Chem. Acc.* **2012**, *131*, 1251.

[22]Ivanic, J. *J. Chem. Phys* **2003**, *119*, 9364.

[23]Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9377.

[24]Docken, K. K.; Hinze, J. *J. Chem. Phys.* **1972**, *57*, 4928.

[25]Lengsfield, B. H., III. *J. Chem. Phys.* **1980**, *73*, 382.

[26]Yarkony, D. R. *Chem. Phys. Lett.* **1981**, *77*, 634.

[27]Fletcher, G. D. *Mol. Phys.* **2007**, *105*, 2971.

[28]Barbatti, M.; Aquino, A. J. A.; Lischka, H. *Mol. Phys.* **2006**, *104*, 1053.

[29]Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.

[30]Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

[31]Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

[32]Wigner, E. *Phys. Rev.* **1932**, *40*, 749.

[33]Butcher, J. *J. Assoc. Comput.* **1965**, *124*.

[34]Granucci, G.; Persico, M. *J. Chem. Phys.* **2007**, *126*, 134114/1.

[35]Hammes-Schiffer, S.; Tully, J. C. *J. Chem. Phys.* **1994**, *101*, 4657.

[36]Pechukas, P. *Phys. Rev.* **1969**, *181*, 174.

[37]Tully, J. C. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1991**, *25*, 299.

[38]Coker, D. F.; Xiao, L. *J. Chem. Phys.* **1995**, *102*, 496.

CHAPTER 7. GENERAL CONCLUSIONS

$O(^3P) + C_2H_4$ Potential Energy Surface:

CASSCF and MRMP2 usually provide accurate barriers and minimum energy paths for the lowest-lying triplet of O + ethylene; and additional single point calculations with more states for the singlet and triplet at the triplet geometries yield preliminary surface crossing data. Overall, the lowest-lying triplet barriers qualitatively agree with Nguyen's[1] barriers. However, not all barriers lie within chemical accuracy of one another. For reactions 1, 3, and 4, additional MR-AQCC calculations give differences from the MRMP2 values of less than 1 kcal/mol. However, some exceptions for these MR-AQCC calculations are: differences of ~2 kcal/mol for reaction 3 Barrier 2 and ~3 kcal/mol for Barrier 1 in reaction 1. In the latter case, near-degenerate states contribute to these larger energy differences and will require other approaches (e.g. SA-MCSCF). Because of the current, typical difficulties in constructing active space sizes smaller than FORS, CASs with and without in-out correlation are constructed. Any differences from these different constructs appear to be minor with respect to energetics. However, placing the lone O2p into the MCSCF core does not always result in a smooth pathway (e.g. reaction 9 Barrier 1).

The application of MRMP2 leads to several items of note. First, as evident from the single point MRMP2 runs, obtaining reasonable energetics for the barriers requires the inclusion of dynamic correlation in some fashion. In this study correlation from MRMP2 lowers energy barriers in many but not all cases; and in most cases, MRMP2 single point energies along the IRC tend to shift the location of energy maxima and minima and hence change the overall barrier. Second, the various CASs lead to similar but not identical barriers that lie within ~1-2 kcal/mol. Third, MRMP2 corrections affect the energetics more in the

cases of bimolecular products to transition state barriers than in the cases of unimolecular intermediate to transition state barriers.

The preliminary surface crossing analysis suggests surface crossing(s) for this system to be around the $\bullet CH_2CH_2O\bullet$ biradical without hyperthermal considerations. In addition, the first and second excited states for the CH_3CHO biradical lie close to one another. As well, in the pathway 3 dissociation, a singlet state lies within ~5 kcal/mol of the lowest-lying triplet transition state. The entire reaction pathway 6 has close-lying triplet and singlet states; and the transition state itself has an additional close-lying second excited state singlet. This nearby state might partially explain why the results from pathway 6 significantly differ from others' reaction 6 results.

$O + C_2H_4$ Excited States and Biradicals:

CASSCF, MRMP2, MR-CISD, and MR-AQCC are used to model the multiconfigurational nature of the singlet and triplet surfaces for the addition of oxygen atom to ethylene. In the biradical region, reaction paths (i.e. paths 1 and 10 – 13) clearly require a multiconfigurational treatment to account for non-dynamical correlation. In particular, with pure state MCSCF solutions (i.e. no SA), TSs 10-13 all have extremely similar geometries with MEPs in a very flat PES region. Relative to available data from single reference treatments, multireference calculations resulted in the following significant difference for TS geometries in this region: for path 20 for one action CH distance, (12,12) CAS shows 1.51 Å whereas Nguyen's[1] B3LYP gives 1.38 Å.

While perturbative, dynamic correlation does not have much effect on singlet stationary point geometries, dynamic correlation is *essential* for obtaining reasonable energetics. Although the MRMP2 captures a significant portion of the dynamic correlation,

this treatment certainly does not encapsulate all of it. Even with multiple levels of theory in this study, the best level of theory for barrier energetics is still not well understood. Since the biradical regions clearly require multiconfigurational wavefunctions, CR-CC(2,3) barriers may require explicit nondynamical correlation while the MRMP2 barriers typically require more dynamic correlation. The MR-AQCC calculations hold promise, but are limited to smaller basis sets and active spaces because of the computational expense. CR-CC(2,3) gives values within 1 kcal/mol of the experimental values for the thermodynamics of pathways 13 and 20 while MRMP2 is within 1.5 kcal/mol for pathway 13 and within about 4.0 kcal/mol for pathway 20. Computational expense limits reaction 20 to a compromised CAS, which may explain a significant part of this error. However, these comparisons with experiment are for mostly closed shell molecules and do not help to elucidate the best methods for obtaining energetics in the biradical regions.

At the MRMP2 and CR-CC(2,3) levels, pathways 1, 10, 11, 12, and M are essentially barrierless, which again emphasizes the need for a careful treatment of this part of the PES. Furthermore, in paths 13 and M, the addition of O2p to the active space qualitatively changes these pathways at the CASSCF level in the sense that geometrically similar TSs lead to completely different sets of minima. However, whenever possible, CASSCF active spaces for pure states do not contain lone O2p orbitals, which tend to lead to IOC ORF with the current, convergence techniques. This active space construction is particularly important in both solving for pure states at the CASSCF level (avoiding IOC ORF) and obtaining single state MRMP2 results (avoiding state root flipping).

In addition to the pure state calculations of the reactions, this study explicitly and appropriately examines the biradical regions. First, two, separate CASSCF surface crossings

between the lowest-lying triplets and lowest-lying singlets are determined. Second, an (8,7) SA-CAS is used to describe the $\bullet CH_2CH_2O\bullet$ biradical around this CASSCF approximation to the given surface crossing. Third, SO-MCQDPT2 is performed on the energetically lowest four states in a region around $\bullet CH_2CH_2O\bullet$. Without treatments such as non-orthogonal MCSCF, these quasi-degenerate states currently require SA wavefunctions to sort out close-lying states – with or without IOC. This active space selection requires care since MCQDPT2 application to an (8,8) SA-CAS reference changes the dominant configuration(s). This fact also supports the avoidance of improper IOC in SA-MCSCF, reference wavefunctions.

Overall, the O + ethylene PES definitely requires a complex treatment in the quasi-degenerate region around the MEXs. In fact, PES regions about paths 10 - 13 will need more sophisticated, multireference methods incorporating dynamic correlation in order to develop an accurate potential for dynamic simulations.

O + C_2H_4 Lowest-lying Singlet:

Selected pathways of the $O(^1D) + C_2H_4$ PES were calculated at the CASSCF, MRMP2, and CR-CC(2,3) levels of theory. Whenever possible, CASSCF active spaces for pure states did not contain lone O2p orbitals, which tend to lead to IOC ORF with the current, convergence techniques. This active space construction is particularly important in both solving for pure states at the CASSCF level (avoiding IOC ORF) and obtaining single state MRMP2 results (avoiding state root flipping). In addition, dynamic correlation proves *essential* for obtaining reasonable energetics. In contrast to the previous triplet study, horizontal shifts in the locations of TSs and minima using single point energetics are small for the singlet reactions in this study. The MRMP2 treatment captures a significant portion

of this correlation but certainly does not encapsulate all of it. In most cases considered here, CR-CC(2,3) results recover additional correlation missed by MRMP2, but in some cases, these coupled cluster results appear to be inappropriate as indicated by the T1 test or further NOCI calculations. Nonetheless, given comparisons of CR-CC(2,3) single point data to both Nguyen[1] and experiment, CR-CC(2,3) optimized barriers appear to be crucial for accurate energetics. Furthermore, CR-CC(2,3) TS optimizations sometimes fail for multiple bond changes in some reactions, for which NOCI expansions reveal more multireference character (i.e. paths 14, 16, 23). These cases and the cases with a possibly insufficient, double zeta basis in the CR-CC(2,3) optimizations limit valid conclusions.

Comparing enthalpies from this work to those available from experiment gives reasonable agreement in many cases. The calculated enthalpy errors range from less than 2.5 kcal/mol to a single, worst case of 6.2 kcal/mol.

ORMAS Nonadiabatic Coupling Calculations:

Based on the results from the $SiCH_4$ and butadiene systems at specific C.I. geometries with carefully chosen CAS and ORMAS active spaces, ORMAS single excitation schemes can often produce CASSCF-like results but tend to have orbital subspace integrity issues, which have the potential to get worse as the subspace sizes increase. On the other hand, schemes with no excitations between subspaces appear to give no orbital subspace issues at the same geometries. As a possible consequence, choosing the smallest, most chemically reasonable orbital subspaces, performing ORMAS with no excitations between those subspaces, and then applying a chosen MR-CI excitation scheme might pedagogically eliminate many orbital issues from calculations. However, for the procedure above, the nondynamical correlation lost should be determined and characterized.

Also based on the same specific systems, this study characterizes the qualitative accuracy of ORMAS NACME in terms of total magnitudes and additionally indicates that CAS and ORMAS NAC vectors strongly coincide at most geometries. The total NACME magnitudes are extremely localized for both CAS and ORMAS, and the largest NACME magnitudes occur close to one another but not at the same geometry. The NACME magnitudes are very sensitive to the proximity of the C.I., and ORMAS gives the appropriate behavior near the CASSCF C.I. Because the ORMAS level of theory can shift the maximum magnitudes, the contours also reveal differences between the CAS and ORMAS total gradient magnitudes (i.e. $\sim 10^{-2} - 10^{-1}$ hartrees/Bohr) about the C.I. geometry with modest energy differences (i.e. 1 – 2 kcal/mol). However, other systems might have several C.I.s in geometrical proximity and thus have more reactivity with 'consecutive', minimum energy C.I.s. Such an idea would complicate the overlaps results. Further investigation is required into C.I.s at the ORMAS level of theory (i.e. C.I.s versus avoided crossings with varying theory level) and how ORMAS NACME affects product distributions during dynamics.

ORMAS Surface Hopping:

The NEWTON-X-GAMESS interface is implemented and tested with the methaniminium cation, CNH_4^+, benchmark system. For this benchmark, the defaults in NEWTON-X were used, and (12,8)CAS results reproduce the characteristic 10 fs S_2 depopulation relative to previous studies[2,3] on CNH_4^+. In addition, ICs and trajectories were run for several active space schemes at the CASSCF and ORMAS levels of theory. With respect to longer trajectory simulation times, the most successful runs contain only the C1s and N1s orbitals in the MCSCF core and experience less energy conservation issues. However, these runs are expensive. MCSCF methods with smaller active spaces and

additional limited excitations into the external space might alleviate some of the energy conservation issues if computational expense does not pose an issue. For the CNH_4^+ test case, the ORMAS approximation with no excitations between orbital subspaces does not reproduce the characteristic 10 fs S_2 depopulation. However, the single excitation ORMAS approximation does produce this depopulation. In addition, different ORMAS schemes might help to identify different state population transfers with individual excitation schemes. In this sense, following different trajectory types with different ORMAS schemes might prove useful in more complicated cases.

References:

[1]Nguyen, T. L.; Vereecken, L.; Hou, X. J.; Nguyen, M. T.; Peeters, J. *J. Phys. Chem. A* **2005**, *109*, 7489.

[2]Barbatti, M.; Granucci, G.; Persico, M.; Ruckenbauer, M.; Vazdar, M.; Eckert-Maksic, M.; Lischka, H. *J. Photochem. Photobiol., A* **2007**, *190*, 228.

[3]Barbatti, M.; Aquino, A. J. A.; Lischka, H. *Mol. Phys.* **2006**, *104*, 1053.